



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/00</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 99/24559</b> <b>(43) International Publication Date:</b> 20 May 1999 (20.05.99)
<b>(21) International Application Number:</b> PCT/GB98/03385 <b>(22) International Filing Date:</b> 11 November 1998 (11.11.98) <b>(30) Priority Data:</b> 9723825.7                      11 November 1997 (11.11.97)    GB <b>(71) Applicant (for all designated States except US):</b> ACTINOVA LIMITED [GB/GB]; 5 Signet Court, Swanns Road, Cambridge CB5 8LA (GB). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> BJORCK, Lars, Henrik [SE/SE]; Magle Stora Kyrkogata 10, S-223 50 Lund (SE). FRICK, Inga-Maria [SE/SE]; Marsvagen 3, S-245 33 Staffanstorps (SE). LEANDERSSON, Tomas, Borje [SE/SE]; Rosenvagen 11, S-245 44 Staffanstorps (SE). AXCRONA, Eugen, Jan, Karol [SE/SE]; Reallinjen 2, S-223 74 Lund (SE). <b>(74) Agent:</b> WOODS, Geoffrey, Corlett; J.A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5LX (GB).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> NUCLEAR TARGETING BY MEANS OF BACTERIAL PROTEINS		
<b>(57) Abstract</b>  A nuclear delivery construct comprises: (i) protein H or a fragment or derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell; and associated therewith (ii) one or more other components whose targeting to the nucleus of the eukaryotic cell is desired.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

-1-

## NUCLEAR TARGETING BY MEANS OF BACTERIAL PROTEINS

### FIELD OF THE INVENTION

5

The present invention relates to nuclear targeting by means of bacterial proteins. It is based on the finding that protein H, which can be derived from *Streptococcus pyogenes*, is targeted to the nuclei of eukaryotic cells, specifically lymphocytes. The present invention therefore relates to nuclear targeting constructs comprising protein H, or a fragment or derivative thereof, associated with another component whose targeting to the nucleus is desired. It also relates to uses of such nuclear targeting of such constructs, and to pharmaceutical compositions comprising them.

10

15

### BACKGROUND TO THE INVENTION

Protein H can be obtained from *Streptococcus pyogenes*, as described in EP-A-0 371, 199 and WO 91/19740. These publications also provide the amino acid sequence of protein H from *Streptococcus pyogenes* and the sequence of the DNA encoding it. Protein H has a characteristic spectrum of immunoglobulin-binding properties, as described in EP-A-0 371,199 and it is also capable of binding albumin (WO 91/19740). In WO 91/19740, a number of regions within protein H were identified, and designated S, A, B, C1, C2, C3 and D regions. Albumin-binding activity was found to be located in the C and/or D regions.

20

25

30

### SUMMARY OF THE INVENTION

-2-

We have now found that protein H has a further, and unexpected, property. When biotinylated protein H was incubated with T-lymphocytes (Jurkat cells), it was found that protein H was targeted to the cell nucleus. We have  
5 also observed the same phenomenon in B-lymphocytes.

In the cell, protein H was found to interact with actin, and with nucleophosmin/B23, a protein known to shuttle between the nucleus and cytoplasm. In the nucleus itself,  
10 protein H was found to interact additionally with the nuclear proteins SET and hnRNP A2/B1, resulting in nuclear accumulation of protein H. We believe that protein H penetrates the cell membrane, becomes associated with nucleophosmin/B23 in the cytoplasm and is  
15 transported across the nuclear membrane into the nucleus, where it interacts with SET and hnRNP. As protein H interacts with actin, the actin cytoskeleton may be involved in transporting protein H from the inside of the cell membrane into the cytoplasm, where it becomes  
20 associated with NPM/B23. However, the possibility that protein H simply diffuses through the cytoplasm cannot be excluded.

The finding that protein H is targeted to the nucleus was  
25 surprising in view of the previously known properties of protein H. Previously, protein H has been known as an immunoglobulin-binding protein, located at the surface of the *Streptococcus* bacterium, where it protects the bacterium by blocking complement activation at the  
30 bacterial cell surface by means of its interaction with the Fc region of IgG. It is therefore surprising that it is also targeted to the cell nucleus, where it may effect a further virulence function on behalf of the *Streptococcus* bacterium. By contrast, protein A, an  
35 immunoglobulin-binding bacterial cell surface molecule

-3-

derived from *Staphylococcus aureus*, showed no nuclear accumulation. Proteins A and H are functionally similar in that both bind to the same site in the Fc region of IgG. The fact that protein H showed nuclear accumulation where protein A did not is therefore all the more surprising.

Based on our findings, protein H can be used to target other molecules to the nucleus. For example, protein H can be used to target drugs to the nucleus, or to target nucleic acids to the nucleus for the purpose of gene therapy.

We have also found that protein H has a cytostatic effect when targeted to the nucleus. In certain situations, for example where a drug is delivered to a proliferating cancer cell for the purposes of treating the cancer, this cytostatic effect will complement the effects of the associated drug. In other situations, it may be desirable to use fragments or derivatives of protein H that do not exhibit the cytostatic effect but are nevertheless capable of being targeted to the nucleus.

The invention accordingly provides a nuclear delivery construct comprising:

- (i) protein H or a fragment or derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell; and
- (ii) one or more other components whose targeting to the nucleus of the eukaryotic cell is desired.

-4-

The invention also provides a pharmaceutical composition comprising such a nuclear delivery construct and a pharmaceutically acceptable carrier.

- 5 The invention also provides such a nuclear delivery construct for use in a method of treatment of the human or animal body, or for use in a method of diagnosis.

10 The invention also provides use of protein H or a fragment or derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell in the manufacture of a medicament for the treatment of the human or animal body in which such a construct is targeted to the nucleus of a eukaryotic cell.

15 The invention also provides use of such a nuclear delivery construct in the manufacture of a medicament for the treatment of disease of the human or animal body in which the construct as defined in claim A is targeted to  
20 the nucleus of a eukaryotic cell.

The invention also provides a method of treating a disease of the human or animal body by targeting such a construct to the nucleus of a eukaryotic cell, said  
25 method comprising administering an effective amount of the construct to the subject.

#### BRIEF DESCRIPTION OF THE DRAWINGS

30 **Figure 1.** Binding of protein H to the surface of human peripheral blood lymphocytes and the human Jurkat T cell line determined by FACS analysis.

-5-

**Figure 2.** Uptake and nuclear accumulation of protein H in Jurkat T cells. Depiction of Jurkat T cells incubated with proteins A or H, cytopinned and stained with FITC-avidin. Protein H (A, B) shows nuclear accumulation, protein A (C) does not.

**Figure 3.** Protein H interacts with nucleophosmin. (A) Results of FPLC on mono-Q column of Jurkat T cells digested with papain and solubilised. (B) Results of radiolabelled pool 85-87 material from (A) being run on protein A-Sepharose (left) and the pooled fractions of the run-through peak from protein A-Sepharose being subjected to a protein H-Sepharose column (right). (C) Identification of NPM in pool 85-87 material from (A). (D) Comparison of SDS PAGE of *in vitro* translated and <sup>35</sup>S-methionine-labelled NPM peptides (left) and SDS-PAGE of the same peptides when applied to protein H-Sepharose. (E) Mapping of the NPM-binding region of protein H by competitive inhibition.

**Figure 4.** Identification of nuclear proteins interacting with protein H. SDS-PAGE gels identifying NPM, protein SET and hnRNP A2/B1.

**Figure 5.** Analysis of the binding of protein H to nucleophosmin and nuclear proteins. (A) Overlay plots of binding of proteins A and H to immobilised NPM (left) or nuclear proteins (right) using plasmon resonance spectroscopy. (B) Competitive inhibition of the binding of <sup>125</sup>I-labelled NPM to protein H-Sepharose with different amounts of unlabeled NPM.

**Figure 6.** Nuclear uptake and cytostatic effect of protein H. (A) SDS-PAGE gels of purified murine B cells incubated

-6-

with LPS and protein H for 24 hours stained with Coomassie (left) or blotted to a PVDF membrane (right), with the blotted membrane probed with an anti-protein H antiserum, followed by peroxidase-conjugated protein A and developed with ECL. (B) Inhibition of proliferation of murine splenic B cells in response to proteins H and A.

**Figure 7.** Schematic representation of protein H.

## 10 DETAILED DESCRIPTION OF THE INVENTION

### Protein H

In WO 91/19740, protein H was characterised as having the following domains, from N-terminal to C-terminal: S, A, B, C1, C2, C3 and D. The S domain is a signal peptide which, in nature, is cleaved from the remaining domains before the mature protein(domains A, B, C1, C2, C3 and D) is translocated to and inserted into the bacterial cell wall.

In the Examples below, a shorter version of protein H is used. This lacks the 41 amino acids of the N-terminal S domain and is also truncated by 30 amino acids at the C-terminus. This version of protein H was produced in *E. Coli*. It is similar to the version of protein H which, in nature, is released from the cell surface of the *Streptococcus* bacterium by the action of a cysteine protease. The version released by the protease also lacks the S domain and is also truncated at the C-terminus. It is slightly shorter than the version produced in *E. Coli*, however.

Herein, unless otherwise stated, the term "protein H"



-7-

means either:

(i) protein H incorporating the signal peptide, as defined in WO 91/19740 and having S, A, B, C1, C2, C3 and D domains and a length of 376 amino acids; or

(ii) mature protein H lacking the S domain and having a length of 335 amino acids; or

(iii) protein H as produced in *E. Coli*, lacking the S domain and truncated by 30 amino acids at the C-terminus, and having a length of 305 amino acids; or

(iv) protein H as cleaved from the *Streptococcus* cell surface by the cysteine protease, lacking the S domain and truncated at the N-terminus by a number of amino acids. The precise length of this version of protein H is not yet known, but its molecular mass suggests that the N-terminal truncation is in the order of 50 amino acids, such that it has a length of approximately 285 amino acids. Thus, this version of protein H may, for example have an N-terminal truncation of from 35 to 45, 45 to 50, 50 to 55 or 55 to 65 amino acids, giving it a length, respectively, of 270 to 280, 280 to 285, 285 to 290 or 290 to 300 amino acids.

The full 376 amino acid sequence is given below (see Sequence information section) as SEQ ID No. 6, together with the coding DNA sequence (SEQ ID No. 5). Amino acid No. 1 represents the beginning of the A domain and the boundaries of each region and each version of protein H are indicated on the sequence. In other words, amino acid No. 1 is the first amino acid of the mature protein (version (ii) of protein H as defined above). The numbering differs from that used in WO 91/19740 in that,

-8-

in WO 91/19740, amino acid No. 1 is the first amino acid of the S domain, which is a signal peptide that is absent from the mature protein.

5 Protein H can be obtained by the methods described in EP-A-0 371, 199 and WO 91/19740, and can also be produced using the methods of Åkesson *et al*, 1990; and Frick *et al*, 1994, in combination with the methods described in the Examples. Any of the above-mentioned versions of  
10 protein H can also be synthesised by recombinant means, based on the sequences given herein and using standard techniques known in the art (as exemplified, for example, by Sambrook *et al*, 1989, Molecular Cloning: A Laboratory Manual). The same applies to fragments and derivatives of  
15 protein H as defined herein. Similarly, protein H and fragments/derivatives thereof can be prepared synthetically by techniques of peptide synthesis already known in the art. This applies especially to fragments of protein H.

20

### **Nuclear targeting**

Based on the experimental results presented herein, protein H is capable of being targeted to the nuclei of  
25 eukaryotic cells. Targeting has been demonstrated in T-lymphocytes (Jurkat cells) and B-lymphocytes (Bjab cells).

For the purposes of the invention, protein H and its  
30 fragments and derivatives are capable of being targeted to the nuclei of eukaryotic cells. A person of skill in the art can determine whether or not any given fragment or derivative is capable of being so targeted in any given cell type using techniques based on those described  
35 in the Examples. For example, the fragment or derivative

-9-

can be biotinylated, then incubated with the cells in question. The intracellular distribution of the biotinylated fragment or derivative can then be determined, for example using FITC-Avidin in conjunction with immunofluorescence microscopy. A person of skill in the art will also be able to devise additional methodologies to determine whether or not protein H or a fragment or derivative thereof is capable of being targeted to the nucleus of any given cell type.

10

Nuclear localisation has been demonstrated in T-lymphocytes (Jurkat cells) and B-lymphocytes (Bjab cells). Preferably, constructs of the invention effect targeting to T-lymphocytes and/or B-lymphocytes. It is not yet clear whether protein H is capable of being targeted to the nuclei of all eukaryotic cell types or whether it is selectively targeted to the nuclei of certain cell types, or whether it is targeted to the nuclei of all cell types to some extent but more strongly in certain cell types. However, this can readily be determined by a person of skill in the art as described above.

15

20

25

30

35

In accordance with the invention, it is preferred that protein H and its fragments/derivatives are capable of interacting with nucleophosmin/B23. It is also preferred that protein H and its fragments/derivatives are capable of interacting with actin. It is also preferred that protein H and its fragments/derivatives are capable of interacting with the nuclear protein SET. It is also preferred that protein H and its fragments/derivatives are capable of interacting with the nuclear protein hnRNP A2/B1. Optionally, protein H and its fragments/derivatives may be capable of interacting with further cytoplasmic and/or nuclear proteins, for example

-10-

other hNRNPs, or transcription factors. A person of skill in the art can determine whether or not a given fragment/derivative does interact with any of the above-mentioned proteins using techniques based on those of the Examples.

Although it is not yet certain whether protein H is capable of being targeted to the nuclei of all eukaryotic cell types or whether it is selectively targeted to the nuclei of some cell types, it is possible to identify some preferred cell types to which protein H and its fragments/derivatives can be used to target associated components in nuclear delivery constructs.

Nucleophosmin (NPM)/B23 is known to act to shuttle proteins between the cytoplasm and the nucleus. It is ubiquitously expressed but it is up-regulated in certain cell types. In particular, NPM/B23 is more abundant in tumour and proliferating cells than in resting cells. It is reasonable to predict that proliferating and tumour cells, having a larger available pool of NPM/B23, will be more effective in translocating protein H and its fragments or derivatives into the nucleus. Thus, according to the invention, it is preferred to use protein H and its fragments/derivatives to target associated components to cells in which NPM/B23 is up-regulated.

Preferred cell types in which protein H and its fragments/derivatives can be used to target associated components to the nucleus include tumour cells, virus-infected cells and healthy but proliferating cells that show increased levels of NPM/B23.

Preferred virus-infected cells are cells infected by the

-11-

human immunodeficiency virus (HIV), for example CD4<sup>+</sup> T-cells; or the human Papilloma virus (HPV), for example cervix epithelial cells and prostate epithelial cells; or a Rhinovirus, for example nasal epithelium cells.

5

Some preferred tumours include rapidly proliferating tumours in general; gliomas and other central nervous system tumours such as neuroblastomas; leukaemias; lymphomas; lung tumours; sarcomas; colon tumours such as carcinomas, e.g. low-grade colon tumours that have shown invasion (Duke III-IV); dispersed renal carcinomas; tubal carcinomas, gastric carcinomas; and prostate carcinomas.

10

Further, constructs of the invention may be used to target therapeutic agents to the nuclei of cells involved inflammatory conditions, as inflammatory conditions commonly involve cell proliferation. Examples of such inflammatory conditions include arthritis, particularly rheumatoid arthritis; arteritis; chondritis; cholangitis; dermatitis; enteritis; myositis; tendosynovitis; and autoimmune inflammatory conditions such as SLE (systemic lupus erythematosus).

15

20

Fibroblasts are a further preferred cell type in which nuclear targeting may be achieved by constructs of the invention.

25

### **The cytostatic effect of protein H**

As described in the Examples, protein H was found to have a cytostatic effect on murine B-lymphocytes, preventing them from proliferating but not inducing apoptosis (cell death). It is not yet clear whether protein H has this effect on all eukaryotic cell types. In situations where protein H and its fragments/derivatives are used to

30

35

-12-

deliver other molecules to the nuclei of proliferating cells, especially tumour cells, this cytostatic effect is likely to be beneficial. In particular, in the case of tumour cells, preventing the cells from proliferating will complement the activity of the associated component and thus assist in combatting the tumour.

In other situations, where it is not desirable to prevent proliferation of the cells, it is preferred to use fragments/derivatives of protein H that are capable of being targeted to the nucleus but which do not exhibit the cytostatic effect. A person of skill in the art will be able to identify such fragments and derivatives by preparing fragments/derivatives by means known in the art and then testing them to determine whether or not they (i) retain the ability to be targeted to the nucleus (see "protein H" above); and (ii) exhibit the cytostatic effect, which can be done using techniques based on those given in the Examples e.g. by exposing the fragments/derivative to proliferating cells and observing whether or not they continue to proliferate.

#### **Fragments and derivatives of protein H**

For the purposes of the invention, a derivative of protein H consists essentially of one of the four amino acid sequences defined herein with respect to SEQ ID No. 6 (see sections entitled "protein H" and "sequence information"). A fragment of protein H is a fragment of any of these four sequences or a fragment of a derivative of any of these four sequences. Such fragments and derivatives are capable of being targeted to the nuclei of eukaryotic cells as described above in the section entitled "nuclear targeting".

-13-

In particular, a derivative of protein H may be an allelic variant or species homologue of protein H which occurs naturally and is capable of being targeted to the nucleus of a eukaryotic cell in a substantially similar manner to the four versions of *Streptococcus pyogenes* protein H as defined herein. Such allelic variants and species homologues will typically be derived from other bacteria, for example, other cocci such as species of *Staphylococcus* or, preferably, *Streptococcus*.

Allelic variants and species homologues can be obtained using techniques known in the art, based on the use of probes derived from the *Streptococcus pyogenes* nucleic acid coding sequence as probes. For example, such a probe can be used to probe libraries made from bacterial cells in order to obtain clones encoding the allelic or species variants. The clones can be manipulated by conventional techniques to express a protein which, according to the invention, is a derivative of protein H.

Preferably, according to the invention, derivatives of protein H have at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% homology with one of the four protein H sequences defined herein with respect to SEQ ID No. 6. More preferably, a derivative will have at least 95%, or at least 99% homology with one of those four sequences over a region of at least 20, preferably at least 30, for instance at least 40, 60 or 100 or more contiguous amino acids. Methods of measuring protein homology are well known in the art and it will be understood by those of skill in the art that in the present context, homology is calculated on the basis of amino acid identity (sometimes referred to as "hard homology").

The sequence of derivatives of the invention may differ

-14-

from that of one of the four protein H sequences defined herein with respect to SEQ ID No. 6 by one or more amino acid substitutions. For example, 1, 2, 3, 4, 5 to 10, 10 to 20 or 20 to 30 substitutions may be present, as long as the derivative has the ability to be targeted to the nucleus of a eukaryotic cell in a substantially similar way to protein H. Preferably, substitutions are conservative. For example, conservative substitutions may be made according to the following table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for one another in a conservative manner.

ALIPHATIC	Non-polar	G A P
		I L V
	Polar-uncharged	C S T M
		N Q
	Polar-charged	D E
		K R
AROMATIC		H F W Y

Similarly, derivatives of the invention may show one or more deletions compared to any of the four protein H sequences defined herein with reference to SEQ ID No. 6. Each deletion may be deletion of, for example, 1, 2, 3, 4 or 5 to 10 amino acids.

Similarly, derivatives of the invention may show one or more insertions compared to any one of the four protein H sequences as defined herein with reference to SEQ ID No. 6. Each insertion may comprise, for example, 1, 2, 3, 4 5 to 10 or 10 to 20 amino acids.

For example, 1, 2, 3, 4 5 or more such deletions or



-15-

insertions may be present.

According to the invention, fragments of protein H, and of derivatives as defined above, may be of any length and may be derived from any region of protein H or one of its derivatives, as long as they have the capacity to be targeted to the nucleus of a eukaryotic cell. For example, suitable fragments may have a length of from 1 to 20 amino acids, from 20 to 50 amino acids, from 50 to 100 amino acids, from 100 to 150 amino acids, from 150 to 200 amino acids, from 200 to 250 amino acids, from 250 to 300 amino acids, from 300 to 350 amino acids, or greater than 350 amino acids.

Preferably, fragments are derived from the more N-terminal regions of protein H. This is because we have found that the A region of protein H, and especially the AB region (i.e. a fragment consisting of the A and B regions) bind to NPM/B23 in a similar manner to complete protein H. We have found that the AB region binds to NPM/B23 more efficiently than complete protein H, and that the A region also binds to protein H, albeit with lower efficiency than complete protein H. The inhibition experiments given in the Examples (see Figure 5C) also suggest that the AB region may be capable of binding to protein SET and hnRNP A2/B1.

This suggests that the A region, and preferably the AB region, may be capable of being targeted to the nucleus. Therefore, fragments of the invention preferably comprise the A region, more preferably the AB region. Optionally, other regions of the protein may also be present.

The AB region of protein H of *Streptococcus pyogenes* is amino acids 1 to 117 in SEQ ID No. 6. Preferably, all of

-16-

these 117 amino acids are present. However, a skilled person will also be able to investigate smaller fragments of the AB region to determine whether or not they retain the capacity to be targeted to the nucleus. In particular, recombinant techniques can be used to generate such fragments and techniques based on those given in the Examples can be used to determine whether or not a given fragment is targeted to the nucleus. Therefore, AB fragments may comprise, for example, 1 to 20, 20 to 50, 50 to 80, 80 to 100 or more of the 117 amino acids of the AB region. In such fragments, the AB region may be truncated at its N-terminus by, for example, 1 to 5, 5 to 10, 5 to 20 or 20 to 50 amino acids and/or at its C-terminus by, for example, 1 to 10, 10 to 20 or 20 to 50 amino acids.

Optionally, derivatives of the invention may comprise protein H or a fragment of protein H as defined herein, and be extended at either the N or the C-terminus or both by an unrelated amino acid sequence. For example, such a sequence may be of up to 10, up to 20, up to 30, up to 50 or up to 100 amino acids in length, or longer (the invention also provides fusion proteins comprising protein H or a derivative or fragment thereof: see below).

In the nuclear targeting constructs of the invention, a revealing label may be present, and may be attached either to protein H or the fragment/derivative of protein H, or to one of the other components. The revealing label may be any suitable label which allows the polypeptide to be detected. Suitable labels include radioisotopes (e.g.  $^{125}\text{I}$ ,  $^{35}\text{S}$ ), an enzyme, an antibody, a polynucleotide or a linker such as biotin.

-17-

In the nuclear targeting constructs of the invention, the protein H component, i.e. protein H or a fragment or derivative thereof, may be present in a purely peptidyl form. Alternatively, it may be chemically modified, e.g. post-translationally modified. For example, it may be glycosylated or comprise modified amino acid residues.

Fragments and derivatives of the invention may be synthesised in any suitable manner. Typically, they will be prepared by recombinant means. However, where appropriate, they may also be made synthetically. For example, they may be made using known techniques of peptide synthesis. This applies especially to fragments of protein H.

Another consideration is the IgG-binding properties of protein H. In general, it is not desirable that fragments/derivatives of protein H retain IgG-binding properties. This is because binding to IgG may lead to the formation of immune complexes which could lead to undesirable side effects and/or reduce protein H's capacity to exert a cytostatic effect. Thus, where the construct of the invention is to be administered by a route that allows the opportunity to form immune complexes (notably intravenous injection), it is preferred that the construct comprises a fragment or derivative that has no capacity to bind to IgG, or at least has a reduced capacity compared to intact protein H. The IgG binding site is around 20 to 30 amino acids long, and spans the boundary between the A and B regions. Thus, it may be desirable to use a derivative of protein H in which this region is deleted or mutated so that it has no capacity, or a reduced capacity, to bind to IgG. Of course, any such modifications should preferably be made without appreciably disrupting the nuclear targeting

-18-

effect of the invention. Thus, it is preferred that such derivatives retain the ability to interact with NPM, SET and hnRNP1.

5 It should also be noted that protein H naturally forms dimers. The dimers have a greater capacity to bind to IgG than isolated protein H monomers. Formation of dimers is favoured below 37°C (i.e. normal human body temperature) but the dimers are less stable above 37°C  
10 (See, for example, Nilson et al Biochemistry, 1995, 34, pp13688-13698). Thus, the IgG binding capacity of protein H *in vivo* may actually be lower than *in vitro* experiments below 37°C suggest. For this reason, even complete  
15 protein H does not necessarily have a great enough IgG binding capacity *in vivo* to disrupt the nuclear targeting effect of the invention.

**Components which may be targeted to the nucleus when associated with protein H and its fragments/derivatives**

20 In principle, the other component of the nuclear delivery construct may be of any nature, as long as its targeting to the nucleus is desired. For example, the other component may be a nucleic acid, i.e. polypeptide, i.e. a  
25 DNA or RNA; a polypeptide, i.e. a protein or peptide, for example an antigen or antibody, e.g. an antibody to a nuclear protein; a small molecule drug; a liposome; or a detectable marker.

30 So far as nucleic acids are concerned, the nuclear targeting properties of protein H represent a novel means of carrying out gene therapy by delivering DNA or RNA to the nucleus. Thus, nucleic acids encoding a protein can be delivered to the nucleus, with a view to their being

-19-

expressed in the nucleus. For example, the nucleic acid may be contained within a nucleic acid construct that effects integration of the nucleic acid into the genome of the cell. Alternatively, the nucleic acid may be  
5 contained within a nucleic acid construct that is designed to replicate within the cell without integration into the genome of the cell, for example within a so-called "mini-chromosome".

10 This avoids the well-known difficulties associated with viral vectors, e.g the risk of uncontrolled viral multiplication.

Protein H and its fragments/derivatives can be used to  
15 target antisense nucleic acids to the nucleus, with a view to their forming a duplex with RNA produced in the nucleus and suppressing the expression of the protein it encodes. The constructs of the invention may thus be associated directly with antisense RNA, or they may be  
20 associated with a construct encoding an antisense RNA, which is transcribed in the nucleus of the cell.

Where the targeting constructs of the invention are used to deliver antisense nucleic acids to cancer (tumour)  
25 cells, antisense RNA can be used to suppress the expression of proteins involved in the cancer phenotype. For example, antisense RNA may be used to suppress the expression of proteins from *myc*, *bcl2*, *bclx* or cyclin genes.

30 Alternatively, antisense RNA to the RNA transcribed from an HLA-DM gene may be used. Nucleophosmin appears to be a common target in chronic graft-versus-host disease after transplantation. Thus, possible modulation of NPM  
35 expression by protein H or derivatives thereof in

-20-

transplanted organs may be developed as a means of reducing the incidence and severity of graft-versus-host disease.

5 HLA-DM is involved in loading peptides on to MHC class II molecules, which leads to their presentation as antigens to CD4 T cells and, ultimately, to the recognition of foreign peptides by T cells. This phenomenon underlies graft-versus-host disease, as recognition of foreign  
10 peptides by T cells is at least partly responsible for the rejection of grafted tissue. In the absence of HLA-DM, only the invariant clip peptide is presented, which does not provoke the host/graft reaction. Therefore, antisense RNA to HLA-DM may be useful in to interrupting  
15 the process by which foreign peptides are presented as antigens by MHC class II molecules, preventing their recognition as antigens by T cells and thus preventing the rejection of grafted tissue by the host tissue and promoting its tolerisatiuon. This potentially applies  
20 both to allografts (where the grafted tissue come from the same species as the recipient host) and to xenografts (where the grafted tissue comes from another species).

So far as small molecule drugs are concerned, one  
25 possibility is to use protein H and fragments/derivatives of protein H to deliver cytotoxic agents. Thus, if protein H exhibits selective nuclear targeting to cells having enhanced B23 expression, particularly tumour cells, protein H represents a way of targeting a  
30 cytotoxic agent to a tumour cell, with a view to killing the tumour cell.

Another possibility is to use protein H and fragments/derivatives of protein H to deliver peptides to  
35 the nucleus. Some preferred peptides are those which

-21-

- chelate certain metal ions, for example  $2+$  ions, such as  $Zn^{2+}$ ,  $Mn^{2+}$  or  $Mg^{2+}$ . Many DNA processing enzymes and transcription factors require  $2+$  ions and limiting the supply of such ions by chelation could reduce the division of cells to which the peptides are targeted. This is expected to be useful in the treatment of tumours, where chelation of certain ions may prevent the proliferation of the tumour cells. For example, targeting of chelating peptides that chelate a particular ion will selectively inhibit certain DNA processing enzymes and transcription factors. If desired, a cytotoxic effect can also be achieved by increasing the amount of chelating peptide used.
- Other preferred peptides are those that interact with transcriptional activators or co-activators involved in regulating tumour cell growth or inflammatory responses, the aim being, respectively, to reduce tumour cell growth or to prevent or reduce the inflammatory response.
- Where protein H or a fragment or derivative thereof is attached to a detectable marker, any suitable marker which allows the construct to be detected may be used. Suitable markers include radioisotopes (e.g.  $^{125}I$ ,  $^{35}S$ ), an enzyme, an antibody, a polynucleotide or a linker such as biotin. In particular, markers that allow detection of the construct by imaging techniques are preferred. Such imaging techniques may be used in diagnostic methods.
- Thus, if protein H is selectively targeted to the nuclei of certain cell types, it could be used as a marker for those cell types. In particular, if protein H is selectively targeted to the nuclei of certain tumour cells, or certain virally infected cells, then protein H or fragments/derivatives thereof can be labelled and used

-22-

as a marker for those cell types. This can be used to diagnose the presence of such cells. Thus, suitably labelled, protein H or fragments/derivatives thereof can be used to diagnose the presence of tumour or virally infected cells to which protein H is selectively targeted. Once such imaging techniques reveal that protein H or a fragment or derivative thereof accumulates in the nuclei of the cells of a given tumour, or type of tumour, a further construct of the invention can also be used to target therapeutic agents to that tumour, or type of tumour.

In the nuclear delivery constructs of the invention, there may be one component which is protein H or a fragment or derivative thereof, as defined herein, and one component which is another molecule (i.e. a 1:1 ratio). Alternatively, the two components may be present with a different ratio, for example 1:2, 1:3, 1:4, 1:5 or higher; or 5:1 or greater, 4:1, 3:1, 2:1 or 1:1. Where there is more than one component associated with protein H, these components may be the same or different. Similarly, there may be more than one molecule of protein H and more than one molecule of the other component, e.g. a ratio of 2:2, 3:3, 2:3 or 3:2.

25

**Association between protein H and its  
fragments/derivatives and the other component(s) of the  
nuclear delivery construct**

30 The other component(s) may be associated with protein H in any manner, as long as the association is strong enough for protein H to transport the other component(s) into the nucleus. The other component(s) may therefore be associated with protein H by any means known in the art.



-23-

For example, the association may be covalent associated. Alternatively, the association may be non-covalent.

For example, the other component(s) may be conjugated to  
5 protein H or a fragment or derivative of protein H. Where  
the other component(s) is/are covalently attached, this  
attachment may be at either the C-terminus or the N-  
terminus, or anywhere between the two termini. In  
constructs where the other component is a polypeptide, it  
10 may be attached by a peptide bond.

Where protein H or a fragment/derivative is non-  
covalently associated with the other component(s), the  
association may be, for example, by means of hydrophobic  
15 interactions, hydrogen bonding or electrostatic  
interactions.

Where the other component is peptidyl in nature; e.g. a  
protein, polypeptide or peptide, it may be desirable to  
20 produce the construct as a fusion protein between protein  
H and the second protein. Thus, a composite DNA sequence  
encoding protein H or a fragment/derivative and the other  
peptidyl component can be prepared by techniques known in  
the art and a fusion protein can be expressed  
25 recombinantly from this composite sequence. Optionally, a  
linker nucleic acid sequence may be provided between the  
two coding sequences, such that a linker region is  
present in the expressed fusion protein. Alternatively,  
the coding sequences of protein H or a  
30 fragment/derivative of protein H may be joined directly  
together in-frame. Fusion proteins are discussed in more  
detail below.

### **Fusion proteins**

-24-

In a fusion protein of the invention, protein H and the other peptidyl component may be present in either orientation, i.e. protein H or a fragment or derivative thereof may be C-terminal or N-terminal to the other peptidyl component. A linker peptide may be present between the two components. Typically, the linker will be flexible, allowing movement of the protein H component with respect to the other peptidyl component. Preferably, the linker will not inhibit the correct expression or folding of either of the two components. Preferably, the linker will not be toxic or immunogenic.

Typically, the peptide linker comprises amino acids that do not have bulky side groups and therefore do not obstruct the folding of the protein component. Further, it is preferred to use uncharged amino acids in the linker. Preferred amino acids for use in linkers include glycine, serine, alanine and threonine.

The peptide linker may be of any suitable length which allows correct folding of the two components. The linker may be from one to four amino acids in length. Alternatively, the linker may be from 5 to 50 amino acids in length, for example 10 to 30 amino acids or 15 to 25 amino acids in length.

Preferably the linker consists essentially of one or more glycine residues and one or more serine residues. Such a linker is termed herein a glycine-serine linker. The linker may contain from 1 to 50, typically 5 to 30 or 10 to 20 glycine residues. The linker may contain from 1 to 50, typically 5 to 30, or from 10 to 20 serine residues.

The linker may consist only of glycine and serine residues. One possible type of glycine-serine linker may

-25-

comprise the sequence (Gly-Gly-Gly-Gly-Ser)<sub>n</sub>. n may be any integer from 2 to 10, preferably from 2 to 5, more preferably 3 or 4, most preferably 3. Other combinations of glycine and serine can also be suitable linkers.

5

The invention also provides nucleic acid sequences, i.e. DNA and RNA sequences, encoding the fusion proteins of the invention. The invention also provides: vectors comprising these nucleic acid sequences; cells containing  
10 such vectors or nucleic acid sequences; and methods of producing fusion proteins of the invention, comprising expressing the nucleic acid sequence encoding the fusion protein in a cell, and recovering the fusion protein thus obtained.

15

A person of skill in the art will be able to generate nucleic acid sequences encoding fusion proteins by techniques known in the art, and will also be able to generate vectors comprising those sequences and transform  
20 or transfect cells with such vectors in order to achieve expression of the fusion protein.

Typically, in a vector, the nucleic acid sequence encoding the fusion protein will be operably linked to a  
25 control sequence capable of providing for the expression of the fusion protein in the host cell. The term "operably linked" refers to a juxtaposition wherein the coding DNA sequence and the control sequence are in a relationship which permits expression of the coding  
30 sequence under the control of the control sequence. Typically, the control sequence will be a promoter.

Optionally, other components may be present in the vector, for example, any of the following: an origin of  
35 replication; an enhancer; a selectable marker gene,

-26-

typically under the control of a promoter; a terminator sequence; or a polyadenylation sequence. Any such components will be positioned such that they are operably linked to the sequence encoding the fusion protein, i.e. in a position such that they exert the desired effect on the coding sequence.

#### **Medical applications of the invention**

10 Nuclear delivery constructs of the invention can be used to deliver components as described above, in association with protein H or a fragment/derivative thereof, to the nucleus. This will be useful in combatting a number of diseases of the human or animal body.

15 In particular, as NPM/B23 is up-regulated in tumour and proliferating cells, nuclear delivery constructs of the invention will be useful in combatting diseases of proliferating cells, or diseases that involve uncontrolled cell proliferation. In particular, nuclear delivery constructs of the invention can be used to deliver components to tumour cells, thereby combatting the tumour.

25 In this connection, it is desirable to use nuclear delivery constructs of the invention in which the component associated with protein H or a fragment/derivative thereof is a tumour-suppressive drug or cytotoxic drug. As indicated above, the cytostatic effect of protein H may complement the effect of these drugs.

Some preferred tumours for such delivery include rapidly proliferating tumours in general; gliomas and other central nervous system tumours such as neuroblastomas;

-27-

leukaemias; lymphomas; lung tumours; sarcomas; colon tumours such as carcinomas, e.g. low-grade colon tumours that have shown invasion (Duke III-IV); dispersed renal carcinomas; tubal carcinomas, gastric carcinomas; and prostate carcinomas.

Similarly, nuclear delivery constructs of the invention can be used to target drugs to virus-infected cells, thus combatting the viral infection. Preferred virus-infected cells are cells infected by the human immunodeficiency virus (HIV), for example CD4<sup>+</sup> T-cells; or the human Papilloma virus (HPV), for example cervix epithelial cells and prostate epithelial cells; or a Rhinovirus, for example nasal epithelium cells. Thus, where a nuclear delivery construct of the invention comprises protein H or a fragment or derivative thereof in association with appropriate antiviral agent, the construct can combat infection by one of these viruses.

Further, as inflammatory conditions often involve cell proliferation, constructs of the invention may be used to target therapeutic agents to the nuclei of cells involved in inflammatory conditions, thus treating these conditions. Examples of such inflammatory conditions include arthritis, particularly rheumatoid arthritis; arteritis; chondritis; cholangitis; dermatitis; enteritis; myositis; tendosynovitis; and autoimmune inflammatory conditions such as SLE (systemic lupus erythematosus).

Also, where the nuclear delivery construct comprises DNA or RNA, gene therapy can be effected. Any suitable gene may be delivered to the nucleus in this way.

Where the targeting constructs of the invention are used to deliver antisense nucleic acids to cancer (tumour)

-28-

cells, antisense RNA can be used to suppress the expression of proteins involved in the cancer phenotype. For example, antisense RNA may be used to suppress the expression of proteins from *myc*, *bcl2*, *bclx* or cyclin genes, thus combatting the cancer concerned. Alternatively, as described above, antisense RNA to the RNA transcribed from an HLA-DM gene may be used, with a view to combatting graft-versus-host disease.

- 10 The use of nuclear targeting constructs of the invention in the treatment of a condition may be combined with other treatments. In particular, where treatment of a tumour is desired, it may be combined with or used in association with other chemotherapeutic or
- 15 chemopreventive agents for providing therapy against tumours. Similarly, it may be combined with the use of other agents for treatment of viral infections or other conditions.
- 20 Nuclear targeting constructs of the invention may be also be used in diagnostic methods. For example, as described above, protein H or a fragment or derivative thereof may be associated with a detectable marker, especially an imaging agent.

25

#### **Pharmaceutical compositions**

- The present invention also provides pharmaceutical compositions comprising a nuclear delivery construct according to the invention and a pharmaceutically acceptable carrier.
- 30

- Any suitable pharmaceutical formulation may be used. For example, suitable formulations may include aqueous or non-aqueous sterile injection solutions, which may
- 35

-29-

contain anti-oxidants, buffers, bacteriostats, bactericidal antibiotics and solutes which render the formulation isotonic with the bodily fluids of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agent and thickening agents. Some preferred formulation ingredients include mannitol or another sugar and/or phosphate-buffered saline (PBS).

It should be understood that in addition to the ingredients particularly mentioned above, the formulations of this invention may include other agents conventional in the art having regard to the type of formulation in question.

#### **Dosage information**

According to the invention, any effective, non-toxic amount of a nuclear deliver construct of the invention may be administered to a patient. The dose of the nuclear delivery construct may be adjusted according to various parameters, for example the nature of the component which is being delivered to the nucleus, the age, weight and condition of the patient to be treated, the mode of administration used, the condition to be treated, the efficiency of the particular nuclear delivery construct being used in transporting the associated component to the nucleus and the required clinical regimen.

As a guide, it appears that  $10^6$  B-lymphocytes accumulate in the region of 1ng to 10ng of protein H in their nuclei.

This suggests that the amount of nuclear delivery construct administered may be such that the dose

-30-

comprises  $10^{-12}$ g to  $10^{-9}$ g of protein H or a fragment or derivative thereof per cell to which the construct is to be delivered, for example  $10^{-12}$ g to  $10^{-8}$ g, e.g. in the region of  $10^{-12}$ g,  $10^{-13}$ g,  $10^{-12}$ g,  $10^{-11}$ g,  $10^{-10}$ g,  $10^{-9}$ g or  $10^{-8}$ g per cell. The amount of protein H or a fragment/derivative thereof thus depends on what cells it is desired to deliver the construct to, and how many of them there are. The total amount of the construct delivered also depends on the size of the component(s) associated with protein H or the fragment/derivative.

For example, a typical dose of a nuclear delivery construct of the invention might contain 1 to  $1000\mu\text{g}$  of protein H or a fragment/derivative thereof, for example 1 to  $10\mu\text{g}$ , 10 to  $100\mu\text{g}$  or 100 to  $1000\mu\text{g}$ .

These dosages are intended only as a guide since a skilled medical practitioner will be able to determine readily the most appropriate dosage for any particular patient and condition.

Similarly, the skilled medical practitioner will be able to determine the appropriate dosage schedule, which will vary according to the factors given above in respect of dosage amounts. However, single doses and multiple doses spread over periods of days, weeks or months are envisaged.

#### **Routes of administration**

Nuclear delivery constructs of the invention can be formulated for clinical administration by mixing them with a pharmaceutically acceptable carrier or diluent, as described above. For example, they can be formulated for



-31-

topical, parenteral, intravenous, intramuscular or transdermal administration. Of course, the route of administration will be tailored to the particular condition to be treated.

5

For example, where complete protein H, or a fragment/derivative comprising the IgG binding site is used in the construct, it may not be desirable to inject the protein or fragment/derivative intravenously, as this may lead to the formulation of immune complexes with IgG. In these situations, other means of administration are preferred, e.g. direct delivery of the construct to the site where it is needed. For example, in the case of a tumour, it is desirable to inject the protein or fragment/derivative directly into the tumour. However, as noted above, even the IgG binding capacity of complete protein H is not necessarily enough to disrupt the nuclear targeting effect of the invention.

## 20 **EXAMPLE**

The following Example illustrates the invention.

### **Summary**

25 Some strains of the human pathogen *Streptococcus pyogenes* express a surface protein called protein H, which is released from the streptococcal surface by a cysteine proteinase produced by the bacteria. Here we find that soluble protein H binds to the surface of lymphocytes and granulocytes. The molecule is taken up by lymphocytes and transported to the nucleus through a previously unknown intracellular pathway. In the cytoplasm, protein H was found to bind to actin whereas when proteins were solubilised from membrane fractions by papain, protein H

30

-32-

was found to interact with nucleophosmin/B23, a protein known to shuttle between the nucleus and the cytoplasm. In the nucleus, protein H is dissociated from nucleophosmin/B23 and instead forms complexes with the nuclear proteins SET and hnRNP A2/B1, resulting in nuclear accumulation of protein H and a cytostatic effect.

### Experimental procedures

#### Bacterial strain, proteins, bacterial expression, *in vitro* translation, labelling of proteins, coupling of proteins to Sepharose

The group A streptococcal strain AP1 of serotype M1 (Åkesson et al., 1994) was used. Recombinant protein H and peptide fragments corresponding to the AB and A regions of protein H have been described (Åkesson et al., 1990; Frick et al., 1994). For the generation of *in vitro* translated NPM, RNA was prepared from Jurkat T cells using RNazol B (Tel-Test Inc., Friendswood, USA) according to the manufacturer's recommendations. cDNA synthesis was performed by incubation of 5 µg RNA with 1x RT buffer (Gibco BRL), 1 mM dNTP, 10 mM DTT, 0.1 mg BSA/ml, 4.5 µM poly dT<sub>18</sub>, 20 U RNase inhibitor (Boehringer Mannheim) and 200 U M-MLV reverse transcriptase (Gibco BRL) in a final volume of 50 µl for 37°C for 1 hour. NPM was PCR amplified using the 5' primer containing a NarI site

5'-GCAGGGCGCCATGGAAGATTCGATGGACAT-3' (SEQ ID No. 1) and the 3' reverse primer

5'-CAGGAATTCTTATTAAAGAGACTTCCTCCACTGCC-3' (SEQ ID No. 2) containing an EcoRI site. For the generation of NPM peptide fragments by *in vitro* translation two additional

-33-

oligonucleotides were used for PCR amplification: The NH<sub>2</sub>-terminal peptide was generated with the reverse primer

5'-CAGGAATTCTTATTAGCTACCCACCTCCAGGG-3' (SEQ ID No. 3)

5 and the primer

5'-TTGATGAAGGTTCCACAGAAAAAGTAAACTTGCTG-3' (SEQ ID

No.4), was used for the COOH-terminal peptide. The PCR products were blunted and cut with EcoRI, whereas the vector pGem-3Z was cut with HincII/EcoRI, and both were  
10 ligated. *In vitro* translation was done using TNT Coupled Reticulocyte Lysate Systems (SDS, Falkenberg, Sweden) according to the manufacturer's recommendations.

Recombinant protein A was purchased from Pharmacia Biotech, Uppsala, Sweden, and actin from porcine heart  
15 was from Sigma (MO, USA). Protein H was labelled with <sup>125</sup>I using the Bolton and Hunter reagent (Amersham, UK).

Protein fractions 85-87 purified with FPLC were concentrated in an Amicon centricon concentrator (Amicon, Inc., Beverly, MA), and labelled with <sup>125</sup>I using the  
20 chloramine T method (Greenwood et al., 1963). <sup>125</sup>I was from Nordion Int. Co. (Canada). Protein H was dialyzed against 0.1 M NaHCO<sub>3</sub> pH 8.3 + 0.5 M NaCl, and coupled to CNBr-activated Sepharose 4B (Pharmacia Biotech) as previously described (Frick et al., 1995).

25

### **Electrophoresis and Western blot analysis**

SDS-PAGE was performed as described by Laemmli (1970), using a total polyacrylamide concentration of 10% or 13.6  
30 % and 3.3% crosslinking. Samples were boiled for 3 minutes in a buffer containing 2% SDS and 5% 2-mercaptoethanol. Gels were fixed with a mixture of 7% acetic acid and 10% ethanol, dried and autoradiographed. Molecular weight markers were from Sigma. Gels were  
35 stained with Coomassie Blue. Protein fractions were

-34-

applied to PVDF membranes (Immobilon, Millipore, Bedford, MA, USA) using a Milliblot-D system (Millipore). Membranes were blocked at room temperature for 1 hour in VBS (10mM veronal, 0.15 M NaCl pH 7.4) containing 0.25% Tween-20 and 0.25% gelatin. After incubation at room temperature for 3 h with radiolabelled protein in VBS containing 0.1% gelatin, the membranes were washed four times with 1.0 M NaCl, 10 mM EDTA, pH 7.7, 0.25% Tween-20 and 0.25% gelatin. The filters were air-dried and autoradiographed at -70°C using Kodak X-Omat AR films and Kodak X-Omat regular intensifying screens.

#### **Cells and preparation of proteins from membrane fractions from the Jurkat cell line**

For flow cytometric analysis human peripheral blood lymphocytes (PBL) from healthy volunteers were depleted from erythrocytes and prepared by Ficoll separation (Pharmacia Biotech). Jurkat cells, a human T cell line, were cultured in RPMI 1640 supplemented with 7.5% FCS and 20mM sodium pyruvate. Membrane preparations were performed in the cold (0-4°C).  $10^9$  Jurkat cells were homogenized in homogenization buffer (0.05 M Tris-HCl pH 7.5, 0.25 M sucrose, 0.005 M  $MgCl_2$ , 0.025 M KCl) followed by centrifugation at 1000 x g for 10 min. The supernatant was further centrifuged at 105000 x g for 45 min. This supernatant was saved and used as a cytoplasmic fraction whereas the pellet obtained was solubilised in 4 ml 0.01 M Tris-HCl pH 8.0. The protein content was determined with the Coomassie protein assay reagent (Pierce, Boule Diagnostics AB, Huddinge, Sweden). After addition of 0.2 mg papain (Sigma) per mg protein solution, and L-cysteine (Sigma) to a final concentration of 4 mM, the mixture was incubated at 37°C for 45 minutes. To terminate the reaction, 4 ml ice-cold 0.01 M Tris-HCl pH 8.0 and

-35-

iodoacetamide (Sigma) to a final concentration of 6 mM were added, followed by centrifugation for 1 hour at 105000 x g. Papain was removed from the supernatant by chromatography on DEAE Sephadex A-50 (Pharmacia Biotech) equilibrated with 20 mM Tris-HCl, pH 8.0. The column was washed with 5 volumes 20 mM Tris-HCl pH 8.0, and the material was eluted with 3 volumes of a 0.5 M NaCl in this buffer followed by dialysis against 20 mM Tris-HCl pH 8.0. For fractionation of the papain digested membrane proteins, the solution was loaded onto an ion-exchange column (Mono-Q, Pharmacia Biotech) mounted on a fluid pressure liquid chromatograph (FPLC, Pharmacia Biotech). Elution was performed with a 70 ml linear salt gradient (from 0 to 1 M NaCl in 20 mM Tris-HCl pH 8.0). Fractions of 0.5 ml were collected. Murine splenic B cells from Balb/c mice were isolated as previously described (Axcrone et al., 1995). Lymphocytes at  $3 \times 10^5$ /ml were plated out in 96 well plates. B cells were activated with LPS (25µg/ml, Difco) and incubated with proteins H and A at indicated concentrations. One µCi [<sup>3</sup>H]thymidine was added per well for the last 4 hours of a 40 hour culture period, cells were harvested and processed for scintillation counting. Values are mean +/- SD of duplicates. For preparation of nuclear and cytoplasmic extracts from murine B cells, cells were activated for 24 hours at  $3.8 \times 10^6$  cells/ml with LPS (25µg/ml) and protein H at 100µg/ml.

#### **Preparation of nuclear extracts from Jurkat T cells and murine B cells**

Nuclei from Jurkat cells were isolated as described by (Mirkovitch et al., 1984). Preparations of nuclei were resuspended in 100 µl of buffer A (50 mM Hepes buffer, 50 mM KCl, 0.1 mM EDTA, 1 mM PMSF, 1 mM DTT, 10% glycerol)

-36-

and additional buffer A was added to a final volume of 162  $\mu$ l. 13  $\mu$ l of 4 M  $(\text{NH}_4)_2\text{SO}_4$  was added to bring the final concentration to 0.3 M. The cells were rocked for 30 minutes and the viscous material was transferred to a 0.2 ml TLA-100 tube (Beckmann), followed by centrifugation at 10<sup>5</sup> rpm for 10 minutes. 125  $\mu$ l of the supernatant was transferred to a second TLA-100 tube, 75  $\mu$ l of 4 M  $(\text{NH}_4)_2\text{SO}_4$  was added to increase the final concentration to 1.5 M. The solution was centrifuged at 50000 rpm for 5 minutes. The supernatant was removed and the pellet was resuspended in 100  $\mu$ l of buffer A. Extracts were used immediately or stored at -80°C.

#### **Affinity chromatography, competitive binding assay, and plasmon resonance spectroscopy**

The <sup>125</sup>I-labelled pooled fractions 85-87 from Mono-Q and *in vitro* translated NPM peptides were applied onto a Protein A-Sepharose column (Pharmacia Biotech) and the flowthrough fractions were run on a protein H-Sepharose column. The column was extensively rinsed with PBSAT (PBSA + 0.05% Tween-20). Bound proteins were eluted with 3 M KSCN and the radioactivity of the fractions was measured in a gamma counter. Fractions were also analysed by SDS-PAGE. Competitive binding assays were performed as reported (Åkerström and Björck, 1989). Binding kinetics were determined by surface plasmon resonance spectroscopy using a BIACORE X system (Biacore AB, Uppsala, Sweden). Actin and nuclear extracts purified on protein H-Sepharose were immobilized on research grade CM5 sensor chips in 10 mM sodium acetate at pH 4.0 and 4.5, respectively, using the amine coupling kit supplied by the manufacturer, whereas biotinylated NPM was coupled to CM5 sensor chips precoupled with avidin (Biacore). All measurements were carried out in PBST. Analyses were

-37-

performed at 25°C and at a flow rate of 10 µl/min. To calculate dissociation and affinity constants, 35 µl of protein H or proteins A, H or actin were applied in serial dilutions (2n; starting at 600 µg/ml). Surfaces were regenerated with 35 µl 1 M KSCN at a flow rate of 10 µl/min. The kinetic data were analysed by the BIAevaluation 2.2 program (Biacore).

#### **Amino acid sequence analysis**

Proteins were separated by SDS-PAGE and stained with Coomassie Blue. Protein bands were excised and digested in-matrix using trypsin. Peptide fragments were separated by reverse-phase HPLC (Vydac 218TP, I.D. 1.6 x 250 mm) and aliquots were analysed by automated Edman degradation using a model 477A sequenator connected to a model 120A on-line PTH-analyser (Applied Biosystems, Weiterstadt, Germany) and by mass analysis using cyano-4-hydroxycinnamic acid as a matrix on a Voyager-DE MALDI-TOF mass spectrometer (Perseptive Biosystems, Wiesbaden, Germany) (Herrmann et al., 1996; Herwald et al., 1996). The Blast network server at the National Center for Biotechnology Information (Altschul et al., 1990) was used for sequence homology searching.

#### **Antibodies, flow cytometrical and fluorescence microscopical analysis**

For flow cytometric analysis of human PBL, mouse anti-human CD3, CD4, HLA-DP/DQ/DR (Becton Dickinson, San Jose, CA, USA) and CD8 antibodies (Dako Patts, Gentofte, Denmark) were used. As a positive and negative control for Jurkat T cells, an isotype control set of a γ1-FITC, γ2-PE and anti-CD45 PerCP (Becton Dickinson) labelled

-38-

antibodies were used, where the  $\gamma 1/\gamma 2$  antibodies were unspecific fluorochrome-conjugated antibodies. Proteins H and A were biotinylated as previously described (Axcrone et al., 1995), and used in conjunction with FITC-coupled avidin (Sigma). Mouse anti-human HLA-DP/DQ/DR antibodies were detected with goat anti-mouse FITC-coupled antibody (Becton Dickinson). FACS-analysis was performed on a Becton Dickinson FACSsort flow cytometer (Becton Dickinson). Each dot blot and histogram represents the analysis of  $10^4$  gated cells. For fluorescence microscopy  $0.5 \times 10^6$  Jurkat cells were incubated with  $20\mu\text{g}$  biotinylated proteins in culture medium for 30 minutes on ice in a flat bottomed 96 well plate. Cells were washed once, followed by continued incubation in a culture chamber at  $37^\circ\text{C}$  for the indicated times. The cells were washed twice in 4 ml PBS, taken up in  $400\mu\text{l}$  PBS, centrifuged on slides in a cytocentrifuge (Shandon, Cytospin 2) at 550 RPM for 3 minutes. After incubation with FITC-coupled Avidin (Sigma), the cells were examined in a fluorescence microscope (Leica Aristoplan) and photographed.

#### **Experimental procedures using anti-protein H or anti-NPM antibodies**

25

##### **Preparation of anti-protein H $\text{F(ab')}_2$ fragments**

Anti-protein H antiserum was applied to a protein G-Sepharose column. The column was extensively washed with PBS and bound IgG was eluted with 0.1 M glycine-HCl pH 2.0. Eluted IgG was dialyzed against acetate buffer pH 4.5 (70 mM  $\text{CH}_3\text{COONa}$  - 50 mM HCl) followed by proteolytic cleavage with pepsin (ratio of protein:pepsin was 100:1) for 21 hours at  $37^\circ\text{C}$ . The reaction was terminated by raising the pH of the solution to 7.5 with 1 M Tris and



-39-

uncleaved IgG was removed by subjecting the material to affinity chromatography using protein G-Sepharose. Unbound material corresponding to polyclonal anti-protein H F(ab')<sub>2</sub> fragments was collected and dialyzed against  
5 PBS. Coupling of F(ab')<sub>2</sub> fragments to Sepharose 4B (Pharmacia Biotech) was performed as recommended by the manufacturer.

#### **Cloning and expression of nucleophosmin (NPM)**

10 For expression in *E. coli* NPM was PCR amplified using the 5' primer containing an EcoRI site 5'-GCAGGAATTCATGGAAGATTTCGATGGACAT-3' (SEQ ID No 7) and the 3' reverse primer 5'-ATAGCGGCCGCTTATTAAAGAGACTTCCTC-3' (SEQ ID No 8) containing a NotI site. The DNA was cloned  
15 into the prokaryotic expression vector pGEX-6p-1 (Pharmacia Biotech) using the EcoRI and NotI sites. Recombinant NPM fused to Glutathione S-transferase (GST) was expressed and purified according to the manufacturers instructions. After purification on Glutathione  
20 Sepharose, the GST tag was cleaved off using PreScission<sup>a</sup> Protease (Pharmacia Biotech). From 1 l of an over night culture, approximately 1 mg pure NPM was achieved. Antibodies against NPM were raised in rabbits.

#### **25 Preparation of cytoplasmic and nuclear extracts from Jurkat T cells and Detroit cells**

Jurkat T cells and Detroit 562 human (carcinoma) pharynx epithelial cells (ATCC CCL 138), incubated with various proteins, were washed five times with PBS, resuspended in  
30 buffer A (10 mM Hepes, 15 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mM EDTA, 1mM PMSF, 1 µg antipain/ml , 0.5 µg leupeptin/ml) and the cell membranes were lysed using 0.2% NP-40. The cytoplasmic fractions were collected after centrifugation at 1000 x g for 10 min and the

-40-

nuclear pellets were washed twice with PBS and resuspended in 100  $\mu$ l of buffer B (50 mM Hepes buffer, 50 mM KCl, 0.1 mM EDTA, 1 mM PMSF, 1 mM DTT, 10% glycerol). Additional buffer B was added to a final volume of 162  $\mu$ l and 13  $\mu$ l of 4 M  $(\text{NH}_4)_2\text{SO}_4$  was added to bring the final concentration to 0.3 M. The resuspended nuclei were rocked for 30 min at 4°C and the viscous material was transferred to a 0.2 ml TLA-100 tube (Beckmann), followed by centrifugation at 350 000 x g for 10 min. The supernatants corresponding to the nuclear fractions were collected and both cytoplasmic and nuclear fractions were subjected to immunoprecipitation.

#### 15 Immunoprecipitation

Extracts prepared from Jurkat cells, incubated with protein H or protein L (150  $\mu$ g) for 16 hours at 37°C, were immunoprecipitated by using polyclonal antibodies against protein H or protein L, 2  $\mu$ l respectively, for 20 hours at 4°C. 40  $\mu$ l protein A-Sepharose was added and incubation was continued for 16 hours at 4°C. Extracts from Detroit 562 cells, incubated with protein H were immunoprecipitated similarly.

Alternatively extracts from Jurkat cells, incubated with protein H (150  $\mu$ g) for various timepoints at 37°C, were 25 precleared with 50  $\mu$ l glycine-Sepharose, for 2 hours at 4°C, followed by immunoprecipitation using polyclonal anti-protein H F(ab')<sub>2</sub>-Sepharose (100  $\mu$ l) for 16 hours at 4°C.

30 The Sepharose pellets were then washed three times with PBS, boiled for 3 min in buffer containing 2% SDS and 5% 2-mercaptoethanol followed by centrifugation at 8000 x g for 5 min. Supernatants were recovered and subjected to SDS-PAGE and Western blot analysis. Membranes were 35 blocked at 37°C in PBST (PBS + 0.05% Tween-20) containing

-41-

5% skim milk and probed with polyclonal antibodies followed by peroxidase-conjugated protein A and developed with ECL.

## 5 Results

### **Protein H interacts with human lymphocytes and granulocytes**

10 Protein H is released from the streptococcal surface through the action of a cysteine proteinase produced by the bacteria (Berge and Björck, 1995). The *E. coli*-produced fragment of protein H used in this study is similar in size to the fragment released by the  
15 streptococcal enzyme, and in the following Examples protein H refers to this COOH-terminally truncated fragment expressed by and purified from *E. coli*. The interaction of protein H with the surface of T cells and granulocytes was analysed with flow cytometry. Human  
20 peripheral blood lymphocytes were incubated with protein H and the majority (>90%) of the CD3<sup>+</sup>, CD4<sup>+</sup> and CD8<sup>+</sup> cells bound protein H (Figure 1A). When the binding of protein H to cells within the granulocyte gate was analysed, protein H also stained these cells brightly (Figure 1A).  
25 Protein H binds to IgGFc (Frick et al., 1994) and previous work has indicated affinity also for human MHCII antigens (Åkesson et al., 1994). The human Jurkat T cell line was therefore chosen for the subsequent experiments as MHC II expression on these cells could be excluded  
30 with flow cytometry. More than 97% of the Jurkat T cells were found to be protein H<sup>+</sup> as compared to the FITC avidin background. Furthermore, as shown in Figure 1B, Jurkat cells were stained by anti-CD45 but not with unspecific  $\gamma$ 1 FITC/ $\gamma$ 2 PE mouse mAbs.

**Protein H is taken up by lymphocytes and accumulated in the nucleus**

5 Biotinylated protein H was incubated with Jurkat cells for different timepoints. Following incubation, cells were cytopinned, fixed, and FITC-coupled avidin was added. As demonstrated by immunofluorescence microscopy (Figure 2) protein H was gradually accumulated in the  
10 nuclei, and after 8 hours 80% of the nuclei showed staining. In contrast, cytoplasmic staining but no labelling of nuclei was detected when the Jurkat cells were incubated with biotinylated protein A (Figure 2C). Like protein H, protein A of *Staphylococcus aureus* is an  
15 IgGFC-binding bacterial surface molecule. Previous work has demonstrated that protein H binds to murine and human B cells (Axcrone et al., 1995) and as in the case of Jurkat cells, protein H was targeted to the nuclei of the human B cell line Bjab, whereas protein A showed no  
20 nuclear accumulation.

**Protein H are taken up by lymphocytes and epithelial cells**

Further studies were carried out using anti protein H antibodies.  $5 \times 10^6$  Jurkat cells were separately  
25 incubated with 150 µg of protein H or protein L for 16 hours. Cytoplasmic and nuclear extracts were prepared and immunoprecipitation was performed using polyclonal antibodies against proteins H and L, respectively, followed by the addition of protein A-Sepharose.  
30 Precipitated materials were run on SDS-PAGE and blotted to PVDF membranes. The membranes were probed with polyclonal antibodies against proteins H and L, respectively, followed by peroxidase-conjugated protein A and developed with ECL. Protein H was taken up by the

-43-

cells and could be detected in both cytoplasmic and nuclear extracts, whereas protein L was not taken up by the cells (not shown).

Epithelial cells (Detroit 562) were also incubated with protein H (150 µg) and extracts prepared from these cells were immunoprecipitated as described above. Protein H was taken up by the Detroit cells, although a lower amount of protein was detected in the nuclear extracts.

#### 10 **Protein H interacts with nucleophosmin/B23 and actin**

To identify proteins interacting with protein H and mediating its uptake, membrane preparations obtained by subcellular fractionation of Jurkat cells were treated with detergent (NP-40). However, no protein H-binding proteins could be detected in this solubilised material before or following purification by ion-exchange chromatography, gel filtration or affinity chromatography on protein H-Sepharose. To release water soluble peptides from Jurkat cell membrane preparations, papain was used. The solubilised peptides were separated by ion-exchange chromatography and fractions were eluted by a linear sodium chloride gradient (Figure 3A). The fractions were applied in slots to PVDF membranes and probed with radiolabelled protein H. Fractions 85-87 reacted with the probe, and they were pooled. A portion (20 ml) was labelled with <sup>125</sup>I and subjected to affinity chromatography on protein A-Sepharose (Figure 3B). The labelled peptides showed no affinity for protein A but when the run-through fractions from protein A-Sepharose were applied to protein H-Sepharose more than 70% of the radioactivity was bound and eluted with 3M KSCN (Figure 3B). When analysed by SDS-PAGE and autoradiography, this material (peak II) contained two major bands of 18 and 54 kDa, respectively. The run-through material (peak I) gave

-44-

rise to a single band with a molecular mass of approximately 16 kDa. Unlabeled pool 85-87 material was now purified on protein H-Sepharose followed by SDS-PAGE. After staining, the 18 and 54 kDa bands (see Figure 3C) were cut out of the gel, digested with trypsin and separated by HPLC. The amino acid sequences shown in Figure 3C could be determined from HPLC peaks and demonstrated that both bands contained nucleophosmin/B23 (NPM), a protein known to shuttle between the cytoplasm and the nucleus (Borer et al., 1989). Monomeric NPM has a molecular mass of 32 kDa but the protein is known to form oligomers (Schmidt-Zachmann et al., 1987; Herrera et al., 1996), including dimers of 70 kDa also under denaturing conditions (see Umekawa et al., 1993 and Figure 3D). Therefore the 54 kDa band probably consists of dimers of NPM fragments generated by papain cleavage whereas the 18 kDa fragments do not form multimers in SDS-PAGE. Intact NPM and two fragments of NPM covering the NH<sub>2</sub>- and COOH-terminal halves of NPM, respectively, were generated by PCR and *in vitro* translation. These <sup>35</sup>S-methionine-labelled peptides were separated by SDS-PAGE followed by autoradiography (Figure 3D, left). As mentioned above, NPM has a tendency to form dimers-oligomers (Schmidt-Zachmann et al., 1987; Umekawa et al., 1993; Herrera et al., 1996). This property is evident for the intact molecule giving rise to bands of 35 and 70 kDa corresponding to monomers and dimers. In case of the NH<sub>2</sub>-terminal fragment with an apparent molecular mass of 18 kDa, a band corresponding to a trimer of 54 kDa is seen, whereas no distinct oligomers are present in lane 3 where the COOH-terminal fragment was run. This is consistent with previous observations demonstrating that the COOH-terminal part of NPM is not essential for oligomerization (Herrera et al., 1996). When the three NPM peptides were subjected to affinity purification on protein H-

-45-

Sepharose, intact NPM and the NH<sub>2</sub>-terminal fragment bound to protein H. The COOH-terminal fragment, however, did not show affinity for protein H-Sepharose (Figure 3D, right). These results and the fact that the 18 kDa NH<sub>2</sub>-terminal papain fragment of NPM shown in Figure 3C, also has affinity for protein H, map the binding of protein H to the NH<sub>2</sub>-terminal part of NPM. As the COOH-terminal region of NPM contains the signals essential for its localization to the nucleolus (Wang et al., 1993), binding of protein H should not interfere with the targeting of NPM.

Using fragments of protein H in competitive binding experiments the region of protein H interacting with NPM was identified. Protein H was immobilized on Sepharose and radiolabelled NPM (pool 85-87) was added. 80-90 percent of the radioactivity was bound to the Sepharose. As demonstrated in Figure 3E this binding was inhibited by unlabeled protein H and by fragments A and AB of protein H, whereas the effect of protein A was at background level also at high concentration. The inhibition with the NH<sub>2</sub>-terminal fragment AB, which is an even more efficient inhibitor than intact protein H, maps the binding of NPM to this region.

The membrane material subjected to papain digestion was obtained by a two-step centrifugation procedure where the supernatant following the final centrifugation at 105000 g represents a cytoplasmic fraction. Also this material was subjected to affinity chromatography on protein H-Sepharose. A dominating band with an apparent molecular mass of 40 kDa was eluted, and NH<sub>2</sub>-terminal amino acid sequencing established that the band was actin.

**Nucleophosmin is co-precipitated with protein H.**

-46-

1 x 10<sup>6</sup> Jurkat cells were incubated with protein H (150 µg) for various timepoints and in order to analyse if nucleophosmin could be coprecipitated with protein H, nuclear extracts were prepared. The extracts were  
5 precleared with glycine-Sepharose and immunoprecipitated using anti-protein H F(ab')<sub>2</sub>-Sepharose. Precipitated materials were run on SDS-PAGE and blotted to a PVDF membrane. The membrane was probed with polyclonal antibodies against recombinant nucleophosmin, followed by  
10 peroxidase-conjugated protein A and developed with ECL. Nucleophosmin could be detected in the nuclear extracts. Thus, nucleophosmin could be coprecipitated using antibodies against protein H demonstrating an *in vivo* interaction between nucleophosmin and protein H.

15 **Identification of nuclear proteins interacting with protein H**

Unlike the homogeneous nuclear staining seen with protein H, anti-NPM antibodies detect a granular accumulation of  
20 NPM in nucleoli (Borer et al., 1989). To investigate whether protein H after the entry into the nucleus, presumably together with NPM, interacts with nuclear proteins, nuclear extracts from Jurkat cells were run on protein H-Sepharose. Several bands were eluted (see  
25 Figure 4, STAIN, lane 2) but when probed with radiolabelled protein H, only three reacted with the probe (Figure 4, BLOT, lane 2). These bands have apparent molecular masses of 39, 42 and 80 kDa, respectively. The 39 and 80 kDa bands (I and III, respectively) were  
30 identified as the SET protein by microsequencing. This protein was initially described as an oncogene product fused to a protein called CAN (v. Lindern et al., 1992). Three tryptic fragments each of the 39 and 80 kDa bands were subjected to NH<sub>2</sub>-terminal sequencing and all  
35 sequences were related to the SET protein (Figure 4,



-47-

lower section). This fact and the molecular mass of the 80 kDa band, suggest that it represents a SET dimer. The amino acid sequence of the 42 kDa band (band II, Figure 4) identified this band to be heterogeneous nuclear  
5 ribonuclear protein (hnRNP) A2/B1, a member of the hnRNP family (Dreyfuss et al., 1993).

**Further analyses and comparison between the interactions  
of protein H with actin, NPM and the nuclear proteins**

10

To analyse the interactions between protein H and the various intracellular proteins in more detail, plasmon resonance spectroscopy was utilized. In these experiments different amounts of protein H were applied and left to  
15 interact with immobilized actin, NPM or with immobilized protein H-binding nuclear proteins, to the level of saturation (Figure 5A). The NPM used in these experiments corresponds to the material shown in Figure 3C, and to create an experimental situation similar to *in vivo*  
20 conditions, a mixture of the nuclear proteins interacting with protein H was used for both plasmon spectroscopy and competitive binding experiments (see Figure 4, STAIN, lane 2).

25 Figure 5A shows typical sensorgrams for the interactions between protein H-actin, protein H-NPM and protein H-nuclear proteins. On the basis of these experiments, dissociation and association rates were calculated, and used to determine association and dissociation constants  
30 (Figure 5B). The data demonstrate that protein H has high affinity for actin but also readily dissociates from the complex, and that protein H has a higher association rate and a considerably slower dissociation rate for the nuclear proteins as compared to NPM. Also, competitive  
35 binding experiments in which NPM and the nuclear proteins

-48-

simultaneously compete for the binding of protein H, showed that unlabeled nuclear proteins more efficiently inhibited the interaction between radiolabelled NPM and protein H-Sepharose, than unlabeled NPM itself . In contrast, actin did not interfere with NPM-protein H binding (Figure 5C) and neither did actin interact with immobilized NPM (Figure 5A, middle section). In none of these experiments did staphylococcal protein A show affinity for actin, NPM or the nuclear proteins (Figure 5 A and C). In summary, the data on the binding kinetics provide an explanation for the release of protein H from its complex with NPM, and its nuclear accumulation as a result of the binding to the nuclear proteins SET and hnRNP A2/B1.

#### **Soluble protein H has a cytostatic effect on murine B cells**

The interaction with the SET protein, a putative transcription factor and oncogene product, and hnRNP A2/B1, a molecule participating in mRNA processing, suggested that protein H could interfere with various cell functions and that metabolically active cells could be particularly sensitive to protein H. It has been shown that protein H interacts also with murine lymphocytes (Axcrone et al., 1995), and the effect of protein H on the proliferation of these cells was therefore investigated. Initially we investigated whether protein H when added to LPS-stimulated murine B cells, was transported to the nucleus, and Western blot experiments demonstrated that this was the case (Figure 6A). There is no indication that protein H is degraded on its way from the exterior of the cell into the nucleus. Thus, the molecular mass of protein H identified in the medium, the

-49-

cytoplasm and the nucleus is very similar if not identical. As a control, an identical PVDF membrane as in Figure 6A was incubated with pre-immune serum followed by peroxidase-conjugated protein A. No signals were  
5 obtained. Following 24 hours of incubation with protein H, the cytoplasm contained 1.4, and the nuclei 2.4ng of protein H per  $10^6$  B cells. Moreover, addition of protein H to the LPS-stimulated B cells, inhibited proliferation measured as [ $^3$ H]Tdr-uptake in a dose dependent manner  
10 (Figure 6B). A 50 percent inhibition was recorded at the highest protein H concentration tested (50  $\mu$ g/ml). As determined by morphology and DNA laddering, protein H did not induce apoptosis in the LPS-stimulated B cells.

#### 15    **Legends to figures**

**Figure 1. Binding of protein H to the surface of human peripheral blood lymphocytes and the human Jurkat T cell line determined by FACS analysis.**

20

(A) Staining of T cells with CD3, CD4, and CD8 versus protein H. Granulocytes were gated out with side scatter/forward scatter and stained with protein H.

(B) Jurkat T cells were stained with protein H and CD45. Background staining is shown with FITC avidin and CD3 on the granulocyte gated cells and with g1-FITC/g2-PE  
25 antibodies on Jurkat cells.

**Figure 2. Uptake and nuclear accumulation of protein H in  
30    Jurkat T cells.**

Jurkat T cells were incubated with biotinylated proteins H or A, cytopinned and stained with FITC coupled avidin. (A) Incubation with protein H for four hours and (B) for

-50-

eight hours. (C) Incubation with protein A for eight hours.

**Figure 3. Protein H interacts with nucleophosmin.**

5

(A) Membrane preparations from Jurkat T cells were digested with papain and the solubilised peptides were subjected to FPLC on a Mono-Q column. Fractions (0.5 ml) eluted with a linear NaCl gradient were analysed for protein H-binding activity in a slot binding assay, and fractions 85-87 reacted with radiolabelled protein H. (B) Radiolabelled pool 85-87 material was run on protein A-Sepharose without showing affinity (left). In the right section the pooled fractions of the run-through peak from protein A-Sepharose were subjected to a protein H-Sepharose column. (C) Peak fractions of peaks I and II from protein H-Sepharose were separated by SDS-PAGE (10% gel), followed by autoradiography. Unlabeled pool 85-87 material was run in parallel. This gel was stained and bands marked by the arrows were cut out. The peptides of the bands were digested with trypsin and tryptic fragments were separated by HPLC. NH<sub>2</sub>-terminal sequences could be determined from one HPLC peak each of the two bands. These sequences established that both bands contained NPM. Numbers indicate amino acid residue positions in the NPM sequences. (D) *In vitro* translated and <sup>35</sup>S-methionine-labelled NPM peptides were separated by SDS-PAGE (13.6% gel). The gel was dried and subjected to autoradiography (left). Lane number corresponds to the peptide run and the peptides are shown schematically in the lower part of the figure. In the COOH-terminal peptide 3, X indicates putative nuclear localization signals. The three radiolabelled peptides were separately applied to protein H-Sepharose. Following extensive washing, bound material was eluted with 3M KSCN, dialyzed

10

15

20

25

30

35

-51-

against PBS, concentrated and run on SDS-PAGE. The gel was dried and autoradiographed (right). (E) Mapping of the NPM-binding region of protein H by competitive inhibition. The binding of  $^{125}\text{I}$ -labelled NPM to protein H immobilized on Sepharose was inhibited with different amounts of unlabeled intact protein H, fragments AB and A of protein H, or protein A.

**Figure 4. Identification of nuclear proteins interacting with protein H.**

A nuclear extract was prepared from Jurkat T cells (lane 1). The extract was precleared with glycine-Sepharose followed by incubation with protein H-Sepharose. After extensive washing proteins bound to the protein H-Sepharose were eluted with 3 M KSCN, dialyzed against PBS and separated by SDS-PAGE (lane 2). Two identical gels (10%) were run simultaneously; one was stained with Coomassie blue (STAIN), one was blotted onto a PVDF membrane and probed with  $^{125}\text{I}$ -labelled protein H (BLOT). Material corresponding to the three bands indicated was submitted to trypsin digestion, HPLC and  $\text{NH}_2$ -terminal sequencing. Three sequences were obtained from each of bands I and III, showing identity to the SET protein. Band II gave rise to two sequences found in hnRNP A2/B1. The sequences are shown in the lower part of the figure, and numbers indicate where homologous residues in the SET and hnRNP A2/B1 are found.

**Figure 5. Analysis of the binding of protein H to nucleophosmin and nuclear proteins.**

(A) Overlay plots of the binding of proteins H and A to immobilized actin (left), NPM (middle) or nuclear

-52-

proteins (right) using plasmon resonance spectroscopy. Increasing concentrations of protein H were applied for 3 min. each during association phase. Dissociation of bound proteins was measured (expressed in resonance units, RU) following injection of buffer alone.

Affinity rates and dissociation constants for the interactions between protein H and actin, immobilized NPM or nuclear proteins are as follows (values are mean  $\pm$  standard deviation from three experiments).

	Actin- Protein H	NPM- Protein H	Nuclear proteins- Protein H
Association rate ( $10^3 \times s^{-1}M^{-1}$ )	2.5 $\pm$ 0.1	6.7 $\pm$ 0.82	1.7 $\pm$ 0.5
Dissociation rate ( $10^5 \times s^{-1}$ )	150 $\pm$ 6.6	190 $\pm$ 2.9	6.7 $\pm$ 0.6
Association constant ( $10^6 \times M^{-1}$ )	1.7 $\pm$ 0.2	3.6 $\pm$ 0.5	26.2 $\pm$ 9.2
Dissociation constant ( $10^{-8} \times M$ )	60.0 $\pm$ 5.7	28.6 $\pm$ 3.7	4.4 $\pm$ 1.7

(B) Competitive inhibition of the binding of  $^{125}I$ -labelled NPM to protein H-Sepharose with different amounts of unlabeled NPM, nuclear proteins, protein A or actin.

**Figure 6. Nuclear uptake and cytostatic effect of protein H.**

-53-

(A)  $7 \times 10^7$  purified murine B cells were incubated with LPS and protein H for 24 hours. Medium (1), cytoplasmic material (2), nuclear extract (3), and protein H (4) were run on SDS-PAGE and stained with Coomassie (left) or blotted to a PVDF membrane (right). The blotted membrane was probed with an anti-protein H antiserum, followed by peroxidase-conjugated protein A and developed with ECL.

(B) Inhibition of proliferation of murine splenic B cells in response to proteins H and A.

10

**Figure 7.** Schematic representation of protein H.

#### Sequence Information

15 Given below are the cDNA (SEQ ID No.5) and amino acid (SEQ ID No.6) sequences of protein H. The boundaries of the S, A, C1, C2, C3 and D regions are marked, as are the N-termini of the alternative forms of protein H.

20 ATG ACT AGA CAA CAA ACC AAG AAA  
Met Thr Arg Gln Gln Thr Lys Lys  
-41  
Signal peptide (Region S)

25 AAT TAT TCA CTA CGG AAA CTA AAA ACC GGT ACG GCT TCA GTA GCC GTT  
Asn Tyr Ser Leu Arg Lys Leu Lys Thr Gly Thr Ala Ser Val Ala Val

GCT TTG ACC GTT TTG GGC GCA GGT TTT GCA AAC CAA ACA ACA GTT AAG  
Ala Leu Thr Val Leu Gly Ala Gly Phe Ala Asn Gln Thr Thr Val Lys

30 GCG|GAA GGG GCT AAA ATT GAT TGG CAA GAA GAG TAT AAA AAG TTA GAC  
Ala|Glu Gly Ala Lys Ile Asp Trp Gln Glu Glu Tyr Lys Lys Leu Asp  
-1 |1

Mature protein, Region A

35 GAA GAT AAT GCT AAA CTT GTT GAG GTT GTT GAA ACC ACA AGT TTG GAA  
Glu Asp Asn Ala Lys Leu Val Glu Val Val Glu Thr Thr Ser Leu Glu

40 AAC GAA AAA CTC AAG AGT GAG AAT GAG GAG AAT AAG AAA AAT TTA GAC  
Asn Glu Lys Leu Lys Ser Glu Asn Glu Glu Asn Lys Lys Asn Leu Asp

-54-

AAA CTT AGC AAA GAA AAT CAA GGA AAG CTC GAA AAA TTG GAG CTT GAC  
Lys Leu Ser Lys Glu Asn Gln Gly Lys Leu Glu Lys Leu Glu Leu Asp

5 TAT CTC AAA AAA TTA GAT CAC GAG CAC AAA GAG CAC CAA AAA GAA CAA  
Tyr Leu Lys Lys Leu Asp His Glu His Lys Glu His Gln Lys Glu Gln

CAA|GAA CAA GAA GAG CGA CAA AAA AAT CAA GAA CAA TTA GAA CGT AAA  
Gln|Glu Gln Glu Glu Arg Gln Lys Asn Gln Glu Gln Leu Glu Arg Lys  
80 |81

10 Region B

TAC CAA CGA GAA GTA GAA AAA CGT TAT CAA GAA CAA CTC CAA AAA CAA  
Tyr Gln Arg Glu Val Glu Lys Arg Tyr Gln Glu Gln Leu Gln Lys Gln

15

CAA CAA TTA GAA ACA GAA|AAG CAA ATC TCA GAA GCT AGT CGT AAG AGC  
Gln Gln Leu Glu Thr Glu|Lys Gln Ile Ser Glu Ala Ser Arg Lys Ser  
117|118

20 Region C1

CTA AGC CGT GAC CTT GAA GCG TCT CGT GCA GCT AAA AAA GAC CTT GAA  
Leu Ser Arg Asp Leu Glu Ala Ser Arg Ala Ala Lys Lys Asp Leu Glu

25 GCT GAG CAC CAA AAA CTT GAA GCT GAG CAC CAA AAA CTT AAA GAA|GAC  
Ala Glu His Gln Lys Leu Glu Ala Glu His Gln Lys Leu Lys Glu|Asp  
158|159

Region C2

30 AAA CAA ATC TCA GAC GCA AGT CGT CAA GGC CTA AGC CGT GAC CTT GAA  
Lys Gln Ile Ser Asp Ala Ser Arg Gln Gly Leu Ser Arg Asp Leu Glu

GCG TCT CGT GCA GCT AAA AAA GAG CTT GAA GCA AAT CAC CAA AAA CTT  
Ala Ser Arg Ala Ala Lys Lys Glu Leu Glu Ala Asn His Gln Lys Leu

35

GAA GCT GAG CAC CAA AAA CTT AAA GAA|GAC AAA CAA ATC TCA GAC GCA  
Glu Ala Glu His Gln Lys Leu Lys Glu|Asp Lys Gln Ile Ser Asp Ala  
200|201

Region C3

40

AGT CGT CAA GGC CTA AGC CGT GAC CTT GAA GCG TCT CGT GCA GCT AAA  
Ser Arg Gln Gly Leu Ser Arg Asp Leu Glu Ala Ser Arg Ala Ala Lys

45 AAA GAG CTT GAA GCA AAT CAC CAA AAA CTT GAA GCA GAA GCA AAA GCA  
Lys Glu Leu Glu Ala Asn His Gln Lys Leu Glu Ala Glu Ala Lys Ala



-55-

CTC AAA GAA|CAA TTA GCG AAA CAA GCT GAA GAA CTT GCA AAA CTA AGA  
 Leu Lys Glu|Gln Leu Ala Lys Gln Ala Glu Glu Leu Ala Lys Leu Arg  
 242|243

## Region D

5

GCT GGA AAA GCA TCA GAC TCA CAA ACC CCT GAT ACA AAA CCA GGA AAC  
 Ala Gly Lys Ala Ser Asp Ser Gln Thr Pro Asp Thr Lys Pro Gly Asn

10

AAA GCT GTT CCA GGT AAA GGT CAA GCA CCA CAA GCA GGT ACA|AAA CCT  
 Lys Ala Val Pro Gly Lys Gly Gln Ala Pro Gln Ala Gly Thr|Lys Pro  
 285|286

Approx C-terminus of protein H cleaved from *S. pyogenes*|

15

AAC CAA AAC AAA GCA CCA ATG AAG GAA ACT AAG AGA CAG TTA CCA TCA  
 Asn Gln Asn Lys Ala Pro Met Lys Glu Thr Lys Arg Gln Leu Pro Ser  
 C-terminus of protein H produced in

20

ACA GGT|GAA ACA GCT AAC CCA TTC TTC ACA GCG GCA GCC CTT ACT GTT  
 Thr Gly|Glu Thr Ala Asn Pro Phe Phe Thr Ala Ala Ala Leu Thr Val  
 305|306  
*E. Coli*|

25

ATG GCA ACA GCT GGA GTA GCA GCA GTT GTA AAA CGC AAA GAA GAA AAC  
 Met Ala Thr Ala Gly Val Ala Ala Val Val Lys Arg Lys Glu Glu Asn  
 335  
 C-terminus of full, mature protein H

## References

- 30 Åkerström, B., and Björck, L. (1989). Protein L: an immunoglobulin light chain-binding bacterial protein. Characterization of binding and physicochemical properties. *J. Biol. Chem.* 264, 19740-19746.
- 35 Åkesson, P., Cooney, J., Kishimoto, F., and Björck, L. (1990). Protein H - a novel IgG binding bacterial protein. *Mol. Immunol.* 27, 523-531.
- 40 Åkesson, P., Schmidt, K.-H., Cooney, J., and Björck, L. (1994). M1 protein and protein H: IgGFc and albumin-

-56-

binding streptococcal surface proteins encoded by adjacent genes. *Biochem. J.* 300, 877-886.

- 5 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.

- 10 Axcróna, K., Björck, L., and Leanderson, T. (1995). Multiple ligand interactions for bacterial immunoglobulin-binding proteins on human and murine cells of the hematopoietic lineage. *Scand. J. Immunol.* 42, 359-367.

- 15 Berge, A., and Björck, L. (1995). Streptococcal cysteine proteinase releases biologically active fragments of streptococcal surface proteins. *J. Biol. Chem.* 270, 9862-9867.

- 20 Borer, R.A., Lehner, C.F., Eppenberger, H.M., and Nigg, E.A. (1989). Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* 56, 379-390.

- 25 Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. *Ann. Rev. Biochem.* 62, 289-321.

- 30 Frick, I.-M., Åkesson, P., Cooney, J., Sjöbring, U., Schmidt, K.-H., Gomi, H., Hattori, S., Tagawa, C., Kishimoto, F., and Björck, L. (1994). Protein H- a surface protein of *Streptococcus pyogenes* with separate binding sites for IgG and albumin. *Mol. Microbiol.* 12, 143-151.

-57-

- Frick, I.-M., Crossin, K.L., Edelman, G.M., and Björck, L. (1995). Protein H- a bacterial surface protein with affinity for both immunoglobulin and fibronectin type III domains. *EMBO J.* 14, 1674-1679.
- 5  
Herrera, J.E., Correia, J.J., Jones, A.E., and Olson, M.O.J. (1996). Sedimentation analyses of the salt- and divalent metal ion-induced oligomerization of nucleolar protein B23. *Biochemistry* 35, 2668-2673.
- 10  
Herrmann, C., Volknandt, W., Wittich, B., Kellner, R., and Zimmermann, H. (1996). The major vault protein (MVP100) is contained in cholinergic nerve terminals of electric ray electric organ. *J. Biol. Chem.* 271, 13908-  
15 13915.
- Herwald, H., Collin, M., Mller-Esterl, W., and Björck, L. (1996). Streptococcal cysteine proteinase releases kinins: a novel virulence mechanism. *J. Exp. Med.* 184,  
20 665-673.
- Herwald, H., Dedio, J., Kellner, R., Loos, M., and Mller-Esterl, W. (1996). Isolation and characterization of the kininogen-binding protein p33 from endothelial  
25 cells. Identity with the gClq receptor. *J. Biol. Chem.* 271, 13040-13047.
- Holm, S.E., Norrby, A., Bergholm, A.M., and Norgren, M. (1992). Aspects of pathogenesis of serious group A  
30 streptococcal infections in Sweden, 1988-1989. *J. Infect. Dis.* 166, 31-37.
- Mirkovitch, J., Mirault, M.E., and Laemmli, U.K. (1984). Organization of the higher-order chromatin loop: specific

-58-

DNA attachment sites on nuclear scaffold. Cell 39, 223-232.

Schmidt-Zachmann, M., Hÿgle-Dörr, B., and Franke, W.W.  
(1987). A constitutive nucleolar protein identified as a  
5 member of the nucleoplasmin family. EMBO J. 6, 1881-1890.

Umekawa, H., Chang, J.-H., Correia, J.J., Wang, D.,  
Wingfield, P.T., and Olson, M.O.J. (1993). Nucleolar  
protein B23: bacterial expression, purification,  
10 oligomerization and secondary structures of two isoforms.  
Cell. Mol. Biol. Res. 39, 635-645.

v. Lindern, M., v. Baal, S., Wiegant, J., Raap, A.,  
Hagemeyer, A., and Grosveld, G. (1992). Can, a putative  
15 oncogene associated with myeloid leukemogenesis, may be  
activated by fusion of its 3' half to different genes:  
characterization of the set gene. Mol. Cell. Biol. 12,  
3346-3355.

20 Wang, D., Umekawa, H., and Olson, M.O.J. (1993).  
Expression and subcellular locations of two forms of  
nucleolar protein B23 in rat tissues and cells. Cell Mol.  
Biol. Res. 39, 33-42.

25 Wang, D., Baumann, A., Szebeni, A., and Olson, M.O.J.  
(1994). The nucleic acid binding activity of nucleolar  
protein B23.1 resides in its carboxyl-terminal end. J.  
Biol. Chem. 269, 30994-30998.

30

-59-

**CLAIMS**

1. A nuclear delivery construct comprising:
  - 5 (i) protein H or a fragment or derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell; and  
associated therewith
  - 10 (ii) one or more other components whose targeting to the nucleus of the eukaryotic cell is desired.
- 15 2. A nuclear delivery construct according to claim 1 wherein the fragment or derivative of protein H comprises the A region of protein H, or a fragment or derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell.
- 20 3. A nuclear delivery construct according to claim 2 wherein the fragment or derivative of protein H comprises the AB region of protein H, or a derivative thereof that is capable of being targeted to the nucleus of a eukaryotic cell.
- 25 4. A nuclear delivery construct according to any one of the preceding claims comprising, as component (i), a fragment or derivative of protein H which is capable of being targeted to the nucleus of a eukaryotic  
30 cell but does not have a cytostatic effect on the cell.
5. A nuclear delivery construct according to any one of the preceding claims wherein component (ii) is a

-60-

polypeptide, nucleic acid, small molecule drug or a liposome.

6. A nuclear delivery construct according to claim 5  
wherein component (ii) is a polypeptide and the  
construct is in the form of a fusion protein.
7. A nuclear delivery construct according to any one of  
claims 1 to 6 wherein components (i) and (ii) are  
covalently or non-covalently associated.
8. A pharmaceutical composition comprising a nuclear  
delivery construct as defined in any one of the  
preceding claims; and a pharmaceutically acceptable  
carrier.
9. A nuclear delivery construct as defined in any one  
of claims 1 to 7 for use in a method of treatment of  
the human or animal body, or for use in a method of  
diagnosis.
10. Use of protein H or a fragment or derivative thereof  
that is capable of being targeted to the nucleus of  
a eukaryotic cell in the manufacture of a medicament  
for the treatment of the human or animal body in  
which a construct as defined in any one of claims 1  
to 7 is targeted to the nucleus of a eukaryotic  
cell.
11. Use of a nuclear delivery construct as defined in  
any one of claims 1 to 7 in the manufacture of a  
medicament for the treatment of disease of the human  
or animal body in which the construct is targeted to  
the nucleus of a eukaryotic cell.

-61-

12. Use according to claim 10 or 11 wherein the construct is targeted to the nucleus of a cell in which nucleophosmin (NPM)/B23 is up-regulated.
- 5 13. Use according to claim 10, 11 or 12 wherein component (i) of the construct interacts with actin, NPM/B23, protein SET or hnRNP A2/B1.
- 10 14. Use according to any one of claims 11 to 13 wherein the cell is a tumour cell or a proliferating cell.
- 15 15. Use according to any one of claims 10 to 14 wherein the cell is a virus-infected cell.
- 15 16. Use according to any one of claims 10 to 15 wherein the cell is a lymphocyte.
- 20 17. Use according to any one of claims 10 to 16 wherein component (ii) of the construct is a nucleic acid encoding a protein, and targeting to the nucleus effects gene therapy of the cell.
- 25 18. A nucleic acid sequence encoding a fusion protein according to claim 6.
19. A vector comprising a nucleic acid sequence according to claim 18.
- 30 20. A cell comprising a nucleic acid sequence according to claim 18 or a vector according to claim 19.
- 35 21. A method of producing a fusion according to claim 6 comprising expressing a nucleic acid sequence according to claim 18 in a cell, and recovering the fusion protein thus obtained.

-62-

22. A method of treating a disease of the human or  
animal body by targeting a construct as defined in  
any one of claims 1 to 7 to the nucleus of a  
eukaryotic cell, comprising administering an  
5 effective amount of the construct to a human or  
animal subject.

23. A method according to claim 18 wherein the disease  
is a tumour or a viral infection.

10



Fig. 1A.

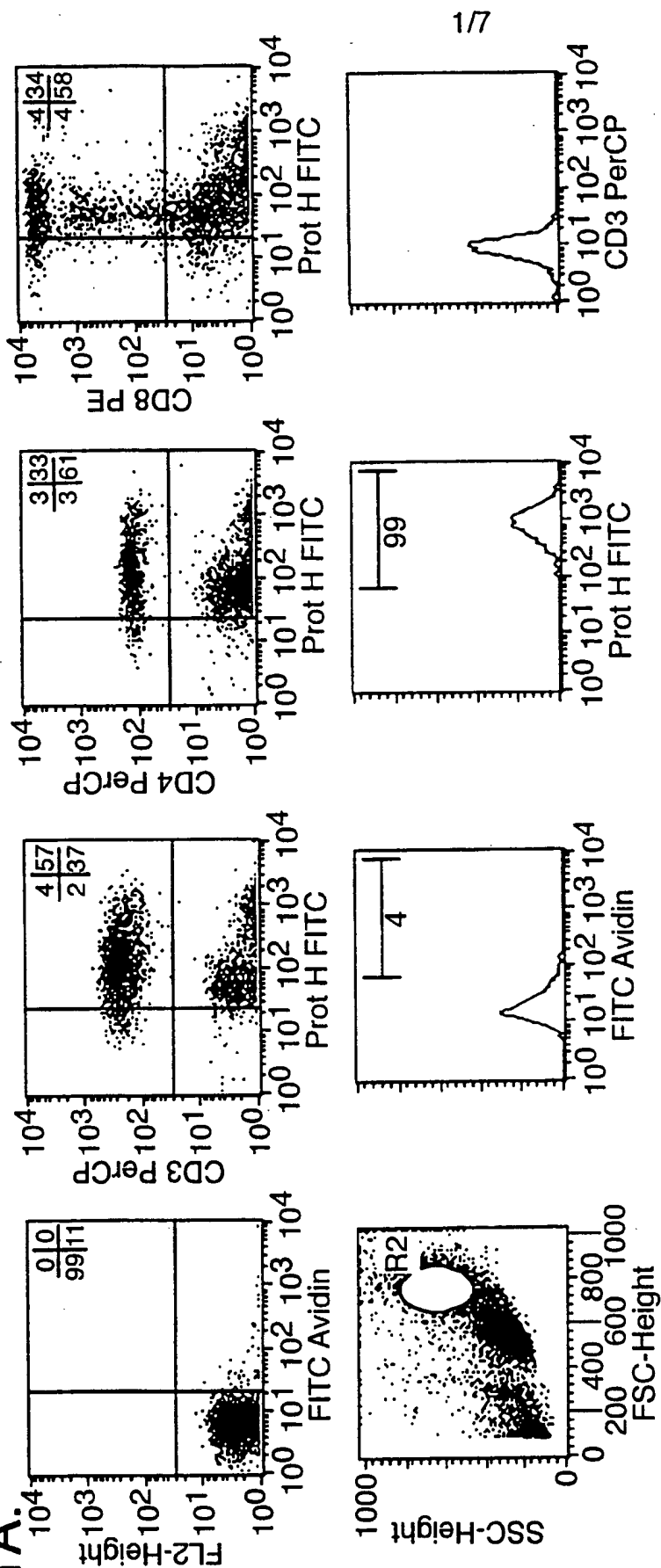
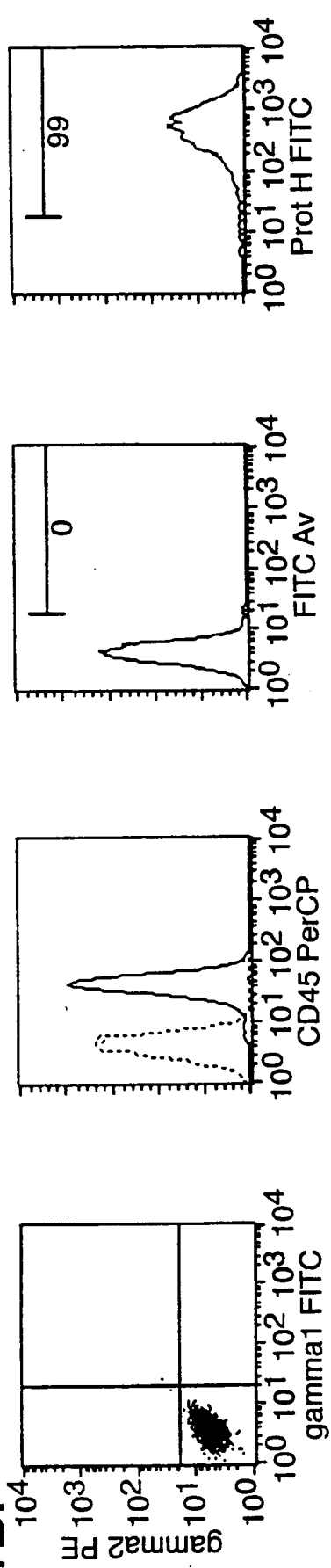


Fig. 1B.



2/7

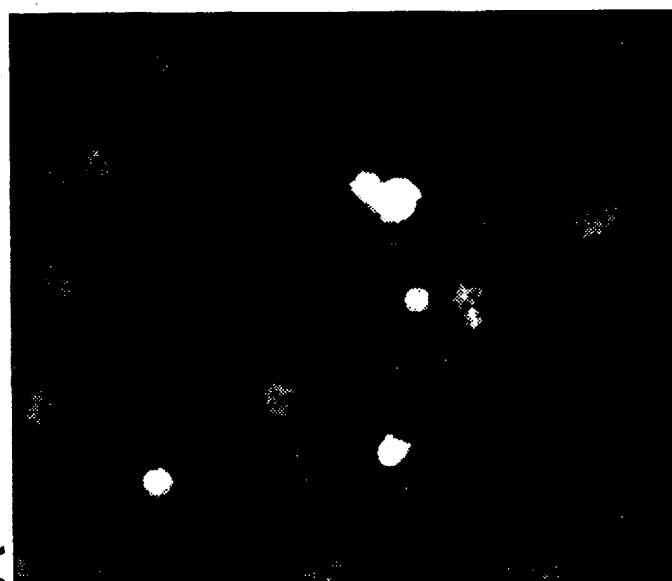
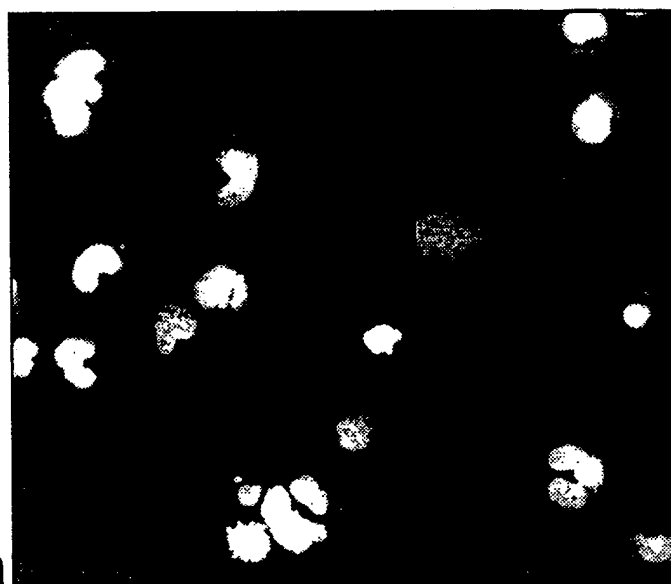
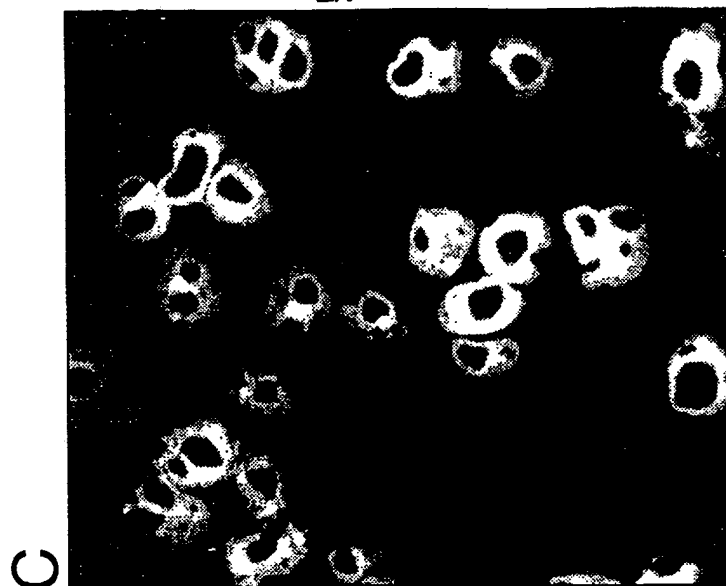


Fig.2.

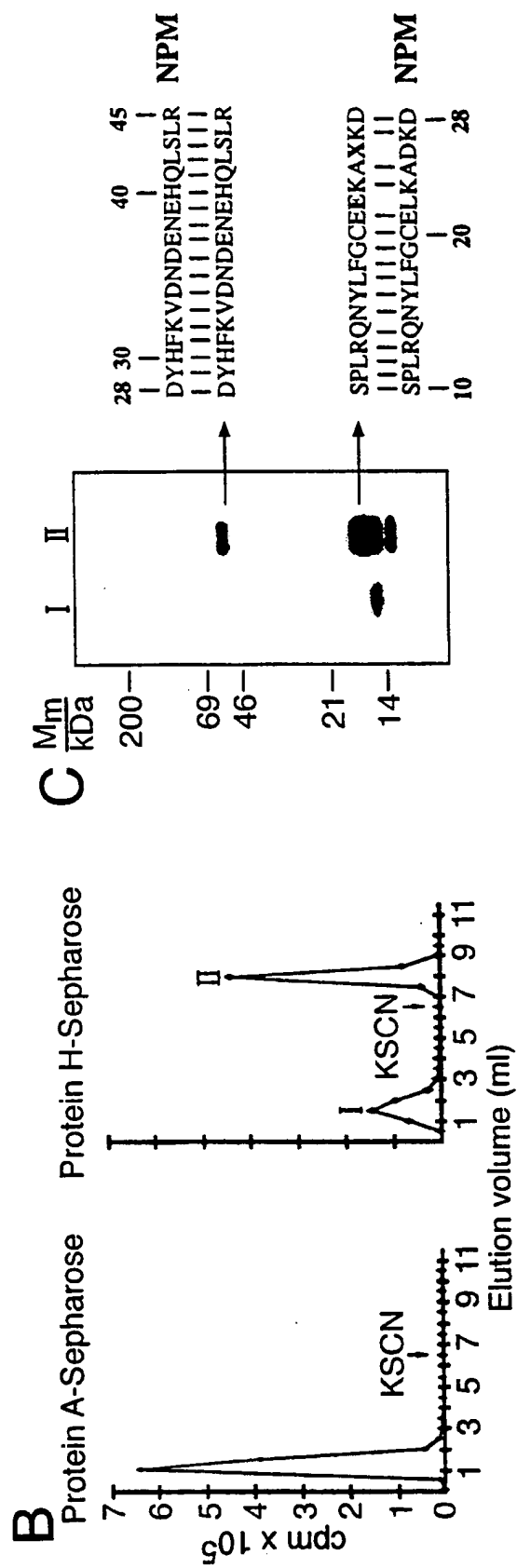
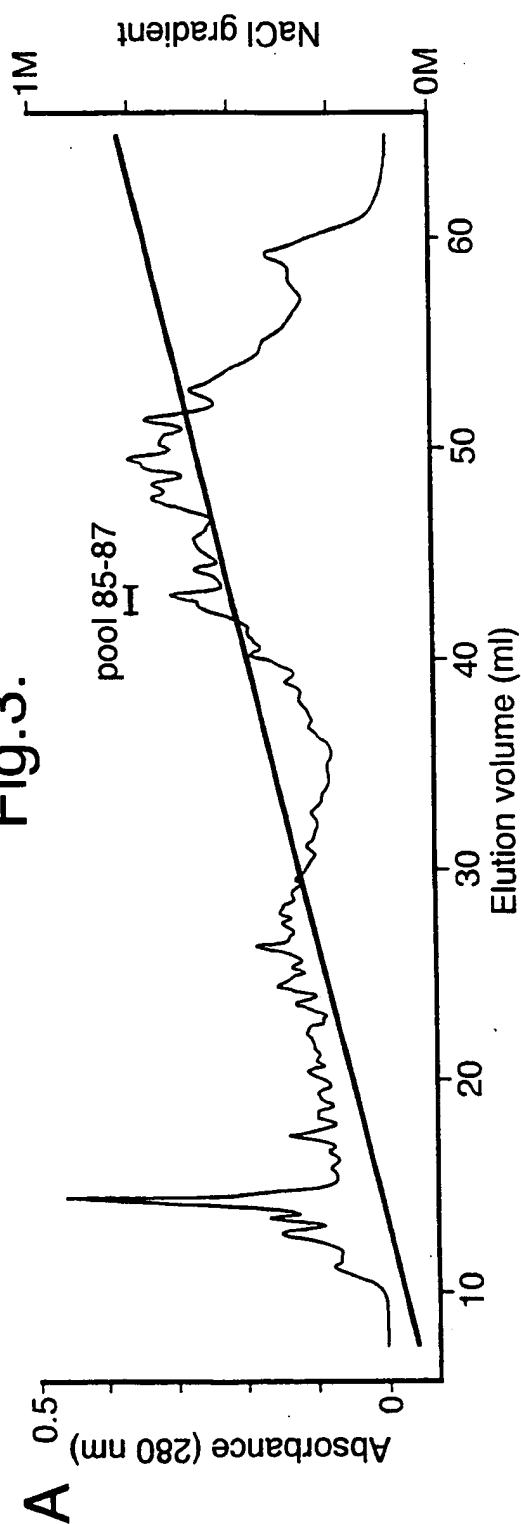


Fig.3 (Cont).

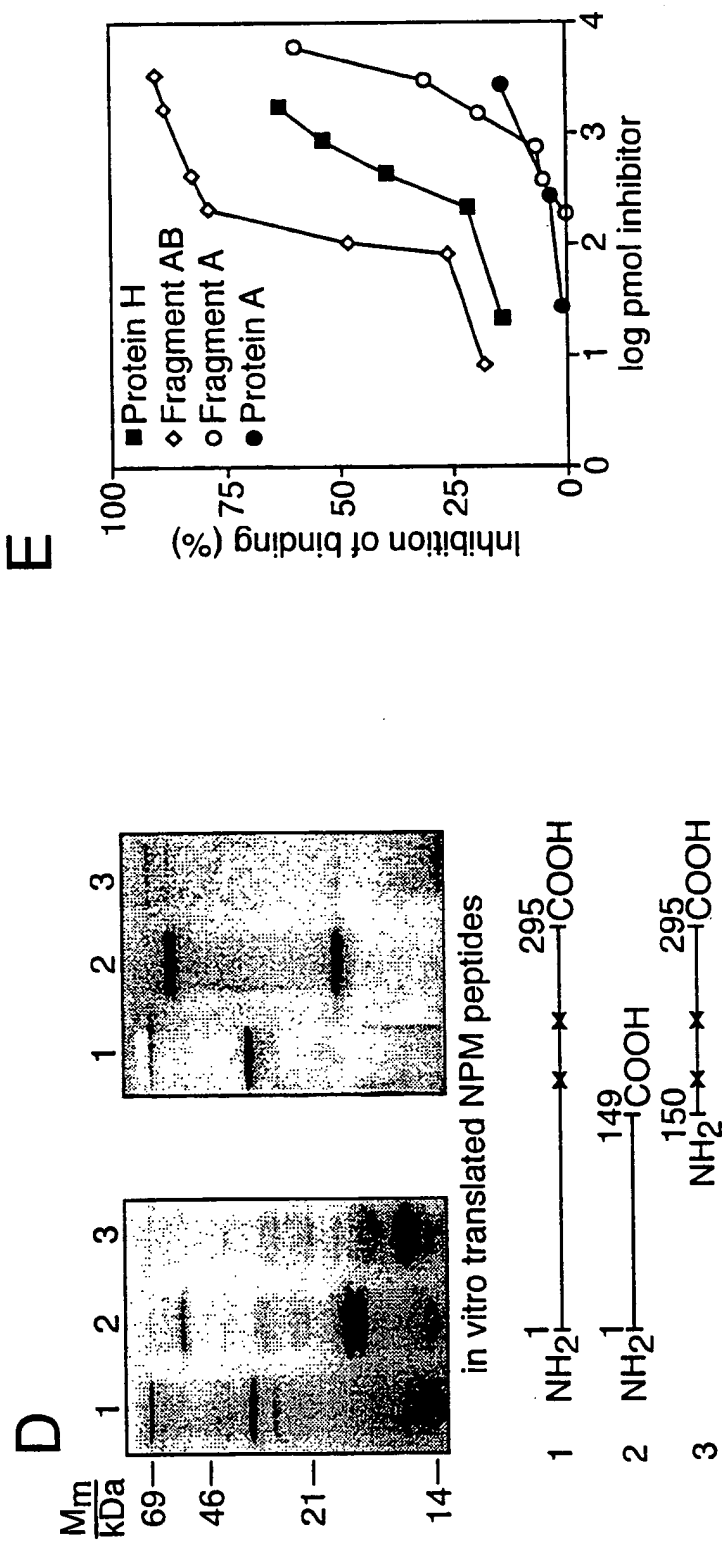
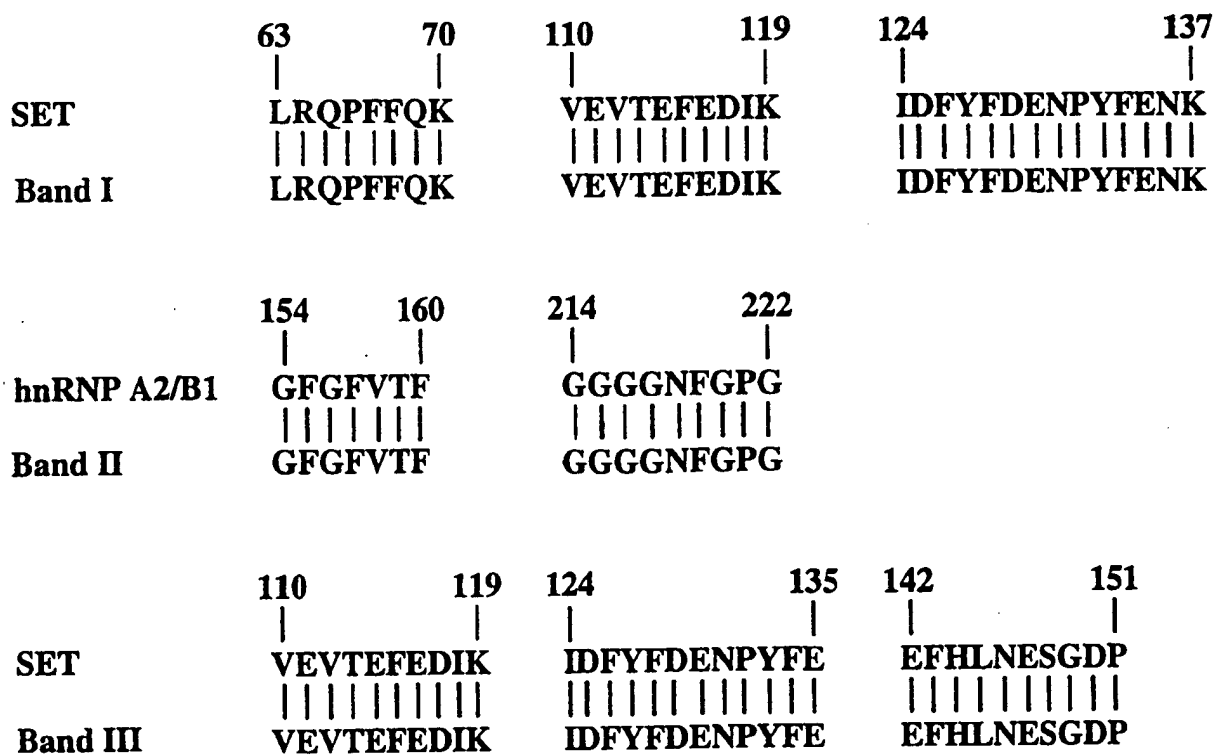
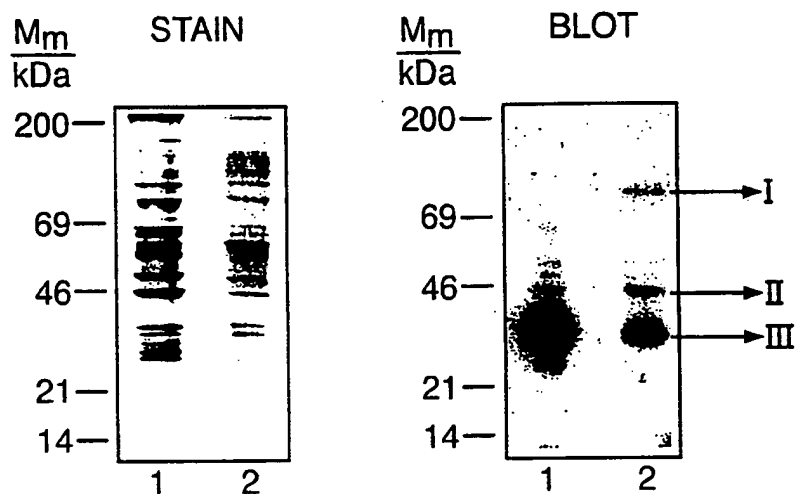


Fig.4.



6/7

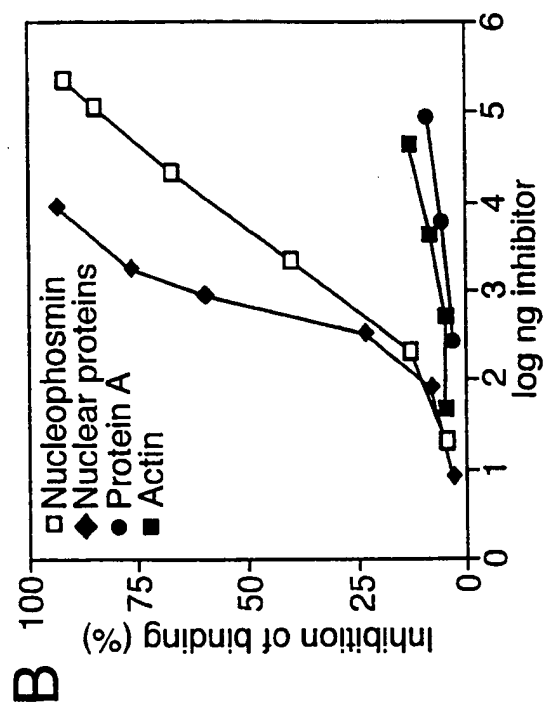
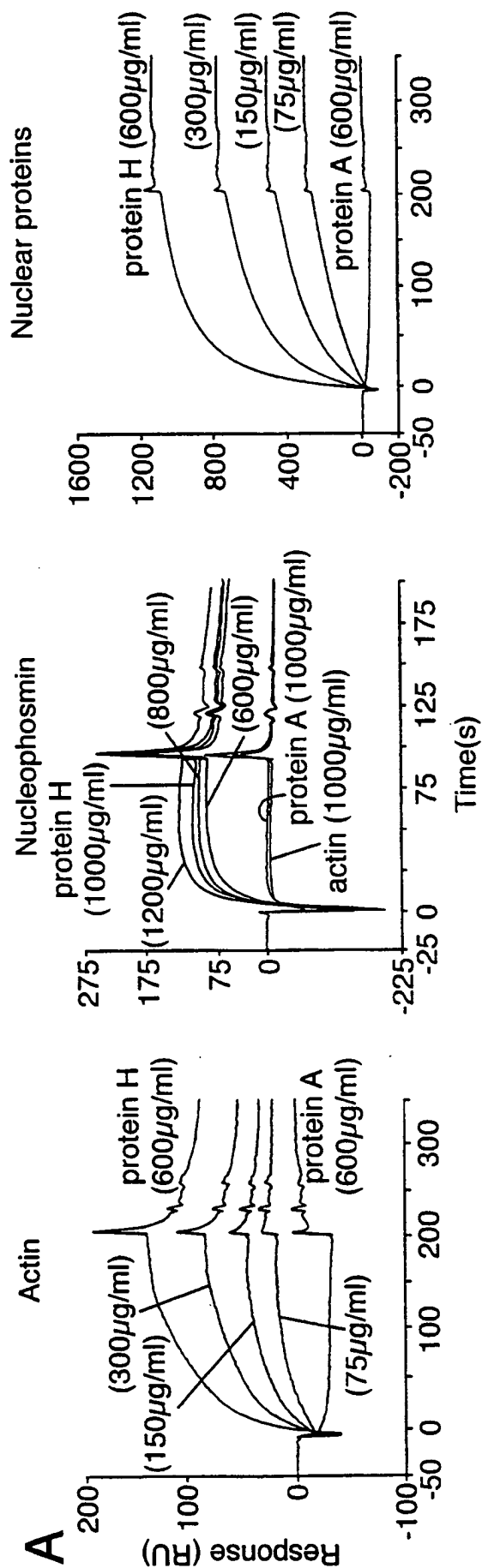


Fig.5.

Fig.6.

